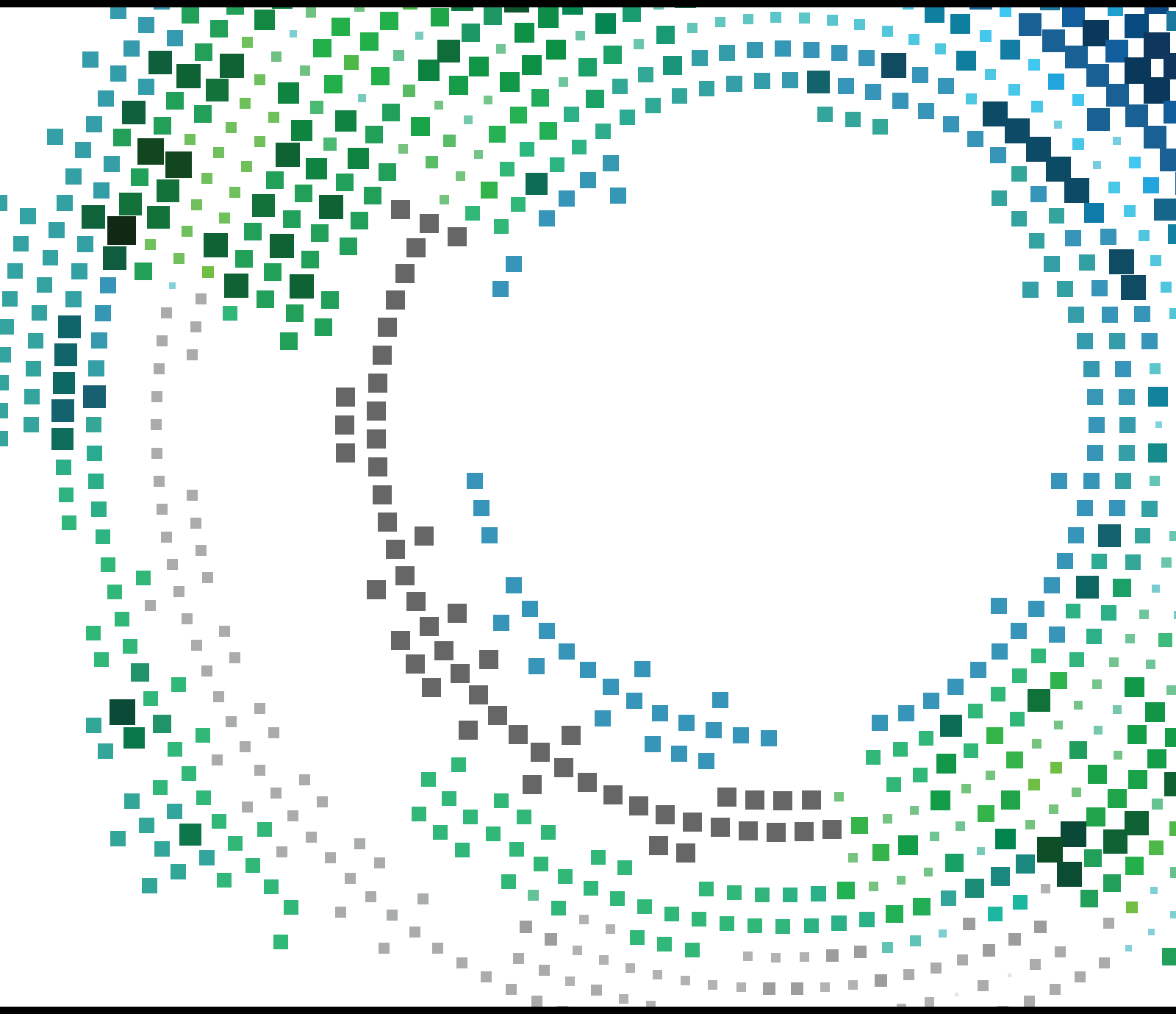
Mobile Information Systems

Ambient Assisted Living and Ambient Intelligence for Health

Lead Guest Editor: Pino Caballero-Gil
Guest Editors: Lilia Georgieva, Ljiljana Brankovic, and Mike Burmester



Ambient Assisted Living and Ambient Intelligence for Health

Mobile Information Systems

Ambient Assisted Living and Ambient Intelligence for Health

Lead Guest Editor: Pino Caballero-Gil
Guest Editors: Lilia Georgieva, Ljiljana Brankovic, and Mike Burmester

Editorial Board

Contents

Hindawi
Mobile Information Systems
Volume 2018, Article ID 7560465, 2 pages
https://doi.org/10.1155/2018/7560465

Editorial

Ambient Assisted Living and Ambient Intelligence for Health

Pino Caballero-Gil,[1] **Lilia Georgieva**,[2] **Ljiljana Brankovic**,[3] **and Mike Burmester**[4]

[1]*Department of Computer Engineering and Systems, University of La Laguna, Tenerife, Spain*
[2]*Department of Computer Science, Heriot-Watt University, Edinburgh, UK*
[3]*School of Electrical Engineering and Computer Science, University of Newcastle, Australia*
[4]*Center for Security & Assurance in IT, Florida State University, USA*

Correspondence should be addressed to Pino Caballero-Gil; pcaballe@ull.es

Received 14 January 2018; Accepted 15 January 2018; Published 20 June 2018

Ambient assisted living is an emerging trend in which artificial intelligence enables the use of new products, services, and processes that help to provide safe, high-quality, and independent lives for the frail and elderly. Due to underlying health issues, aspects of everyday living can become physically and mentally challenging for them. Technology can support daily interaction and be integrated in the health care of senior citizens, which are both vital to ensure their health and happiness.

Artificial intelligence has enabled significant advancements in ensuring such support, while preserving independence. Advancements include development of information and communication technologies used in versatile ways, including for prediction, prevention, rehabilitation, and support. However, technology that enables ambient assisted living comes with its own challenges. It needs to be easy to use, while suitably designed, and adaptable to changing needs and individual preferences.

The paper by X. Ferre et al. addresses the use of ultrasonic sensor-based gait speed measurement device controlled via a mobile interface, which permits patients to self-assess physical performance. This allows for timely detection of functional decline and frailty, which can, if undetected, ultimately progress to disability.

The paper by C. Ramírez-Fernández et al. presents the usability evaluation of a haptic-enhanced tele rehabilitation system for massage therapy of the back. The system includes features that allow for administering online therapy programs, providing self-adjustable and safety treatment of back massages using a virtual environment, and saving and replaying massage sessions according to a patient's therapy program.

The paper by D. Spoladore et al. presents a smart home simulator, using semantic and virtual reality-enhanced configuration of domestic environments, and taking into account both the preferences of the end-users, the configurations of smart appliances, and relevant technologies, including deployment and data-sharing issues.

The paper by I. Rodríguez et al. addresses issues surrounding using mobile and wearable devices for self-reporting of chronic pain and pain management in older adults.

The paper by M. Simón et al. introduces a system for gathering physiological data, which is valuable for the analysis of personal characteristics, such as behaviour, health conditions, and preferences.

The paper by M. A. Teruel et al. discusses physical and cognitive rehabilitation and shows that the development of collaborative rehabilitation systems is one of the best alternatives to mitigate isolation.

The article by M. Espinilla et al. introduces a fuzzy intelligent system for patients with preeclampsia in wearable devices. The system uses a decision analysis tool for the early detection of the condition in women at risk.

Gait analysis, using computer vision based on cloud platform and mobile device, is the topic of the paper by M. Nieto-Hidalgo et al. Since deterioration of cognitive and motor function is linked to gait patterns, gait analysis can be a powerful tool to assess frailty and senility syndromes.

The paper by B. Liu et al. discusses the relevance of monitoring breathing and establishing accurate breathing

rate using a deep learning-based fine-grained breathing rate monitoring algorithm, which works on smartphone and achieves professional-level accuracy.

A fuzzy logic-based personalized method to classify perceived exertion in workplaces using a wearable heart rate sensor is the topic of the paper by P. Pancardo et al. Wearable heart rate sensors represent an effective way to capture perceived exertion, ergonomic methods are generic, and they do not consider the diffuse nature of the ranges that classify the efforts. The proposed method is personalized, and it assesses perceived exertion and uses fuzzy logic as an option to manage imprecision and uncertainty in involved variables.

The paper by D. Martín et al. discusses approaches to improving learning tasks for mentally handicapped people using ambient intelligence techniques based on cyber-physical systems. The paper shows that that such solutions are feasible and allow for learning of complex tasks in some cases.

Pino Caballero-Gil
Lilia Georgieva
Ljiljana Brankovic
Mike Burmester

Hindawi
Mobile Information Systems
Volume 2018, Article ID 5214067, 9 pages
https://doi.org/10.1155/2018/5214067

Research Article

Deep Learning versus Professional Healthcare Equipment: A Fine-Grained Breathing Rate Monitoring Model

Bang Liu, Xili Dai, Haigang Gong, Zihao Guo, Nianbo Liu, Xiaomin Wang, and Ming Liu

Big Data Research Center, Department of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China

Correspondence should be addressed to Xili Dai; daixili_cs@163.com

Received 28 July 2017; Revised 13 November 2017; Accepted 28 November 2017; Published 1 March 2018

Academic Editor: Pino Caballero-Gil

In mHealth field, accurate breathing rate monitoring technique has benefited a broad array of healthcare-related applications. Many approaches try to use smartphone or wearable device with fine-grained monitoring algorithm to accomplish the task, which can only be done by professional medical equipment before. However, such schemes usually result in bad performance in comparison to professional medical equipment. In this paper, we propose DeepFilter, a deep learning-based fine-grained breathing rate monitoring algorithm that works on smartphone and achieves professional-level accuracy. DeepFilter is a bi-directional recurrent neural network (RNN) stacked with convolutional layers and speeded up by batch normalization. Moreover, we collect 16.17 GB breathing sound recording data of 248 hours from 109 and another 10 volunteers to train and test our model, respectively. The results show a reasonably good accuracy of breathing rate monitoring.

1. Introduction

The emergence of mHealth draws much attention both in industry and academy [1]. Google, Microsoft, and Apple conduct a series of work on mHealth from hardware to software. Google is the first one to get involved in mHealth. In April 2012, Google released Google Glass [2] and applied it to healthcare in July 2013 [3]. Pristine declared to develop medical application for Google Glass. After that, Google accomplished the acquisition of a biotech company Lift Labs, which invented an electronic spoon to help Parkinson patients have food. In 2015, Google X announced that it was working on wearable suits which can exam cancer cell of users. In addition, Microsoft Band, Apple Watch, Fitbit, Jawbone, and more smart wearable devices bloom up everywhere.

There exists a broad array of healthcare-related applications on sleep monitoring by smart wearable devices [4]. They often aim at fine-grained breathing rate monitoring as a kind of nonobtrusive sleep monitoring for the understanding of users' sleep quality. Since inadequate and irregular sleep can lead to serious health problems such as fatigue, depression, cardiovascular disease, and anxiety [5], breathing rate monitoring is critical to detect early signs of several diseases such as diabetes and heart disease [6]. The breathing rate monitoring can also be applied to the sleep apnea diagnosis and treatment, treatment for asthma [7], and sleep stage detection [8]. Thus, fine-grained breathing rate monitoring is important to facilitate these healthcare-related applications.

Traditionally, one's breathing rate can be captured by professional medical equipment as monitoring machines in hospitals. In most cases, such machines are too expensive, too complex, and too heavy for daily use for ordinary people. A possible solution is to achieve accurate sleep monitoring via smartphone or other devices with recognition algorithm [9], which is more and more popular in current healthcare-related applications. For example, Ren et al. [10] exploit the readily available smartphone earphone placed close to the user to reliably capture the human breathing sound. It cannot work if the earphone is apart from the user. Liu et al. [11] tracks the vital signs of both the breathing rate and the heart rate during sleep, by using off-the-shelf WiFi without

any wearable or dedicated devices. However, the wearable devices cannot achieve approximative performance in comparison to professional medical equipment. The latter often has a much lower signal-to-noise ratio (SNR) and can achieve a much higher accuracy in breathing rate monitoring.

In this paper, we aim at developing a fine-grained breathing rate monitoring algorithm that works on smartphone and achieves professional-level accuracy. We propose a deep learning model such as DeepFilter, which can filter the breathing from low SNR data. We empirically exploit the framework of deep learning and apply it to the fine-grained breathing rate monitoring on smartphone. The deep learning model combines several convolutional layers and a bidirectional recurrent layer and is trained in an end-to-end manner using the cross entropy loss function. In addition, batch normalization is adapted to speedup the training. Moreover, we collect 16.17 GB breathing sound recording data of 248 hours from 109 and another 10 volunteers to train and test our model, respectively. The results show a reasonably good accuracy of breathing rate monitoring.

The main contributions of this paper are highlighted as follows:

(i) As to our best knowledge, we are the first to apply deep learning to fine-grained breathing rate monitoring, with low SNR data recognition.

(ii) We run real experiments on smartphone and verify the availability and performance of our model, which directly promotes the sleep monitoring applications in our daily lives.

2. Related Work

Since our scheme involves accurate sleep monitoring and deep learning, we mainly discuss the previous work on the two aspects.

Medical-based sleep-monitoring systems are often developed for clinical usage. In particular, polysomnography [12] is used in medical facilities to perform accurate sleep monitoring by attaching multiple sensors on patients, which requires professional installation and maintenance. It can measure many body functions during sleep, including breathing functions, eye movements, heart rhythm, and muscle activity. Such systems incur high cost and are usually limited to clinical usage. DoppleSleep [13] is a contactless sleep sensing system that continuously and unobtrusively tracks sleep quality, by using commercial off-the-shelf radar modules.

Some smartphone apps, such as Sleep as Android, Sleep Cycle Alarm Clock, and iSleep [14], can perform low-cost sleep monitoring by using the smartphone built-in microphone and motion sensors. These apps, however, only support coarse-grained monitoring, such as the detection of body movements, coughing, and snoring [15], and utilizes the phone usage features such as the duration of phone lock to measure sleep duration. The Respiratory app [16] derives a person's respiratory rate by analyzing the movements of the users' abdomen when placing the phone between the users' rib cage and stomach. ApneaApp [17] is a contactless sleep apnea event detection system that works on smartphone, which does this by transforming the phone into an active sonar system that emits frequency-modulated sound signals and listens to their reflections to capture the breathing movement. Ren et al. [10] exploit the readily available smartphone earphone placed close to the user to reliably capture the human breathing sound. Liu et al. [11] propose to track the vital signs of both the breathing rate and the heart rate during sleep by using off-the-shelf WiFi without any wearable or dedicated devices. There is still a gap between the performance of professional equipment and that of the approaches above.

Recently, deep neural networks are first used for better phone recognition [18], in which traditional Gaussian mixture models are replaced by deep neural networks that contain many layers of features and a very large number of parameters. Convolutional networks have also been found beneficial for acoustic models. Recurrent neural networks are beginning to be deployed in state-of-the-art recognizers [19] and work well with convolutional layers for the feature extraction [20]. We are inspired by the good performance of the previous work on speech recognition and introduce deep learning algorithm into the problem of fine-grained breathing rate monitoring [21].

3. DeepFilter

In this section, we introduce the whole framework of DeepFilter and investigate the training of the model in detail.

3.1. The Framework. Figure 1 shows the framework of DeepFilter. Our model is a RNN that begins with several convolutional input layers, followed by a fully connected layer and multiple recurrent (uni- or bidirectional) layers, and ends with an output layer. The network is trained end to end and is added batch normalization with cross entropy loss function.

In speech recognition, "end-to-end" is often used to support the training without aligned training labels, which does not involve the frame-level cross entropy loss function. In breathing rate monitoring, we first create frame-aligned training labels and translate them into a classification problem to decide whether the frames belong to inhaling/exhaling or not. Since one inhaling/exhaling event may involve several frames, we exploit recurrent layers to process input sequence, for the recurrent layers can capture the sequence information to improve the performance. Thus, the input/output sequence of our model is similar to that of speech recognition. The only difference is that the input and output sequences of speech recognition have different lengths while that of our model have the same length.

To one sample x^i and label y^i, sequence frames are sampled from training set $\chi = \{X^1, X^2, \ldots, X^N\}$, which generates some voice recordings of size N. We assume that X^i is a sound recording for 136 seconds (the sampling rate is 44100 Hz), and the duration of each frame is 40 ms (which is an empirical value that always used as window size in speech recognition) It can be divided into samples $X^i = \{x^1, x^2, \ldots, x^n\}$, and each sample $x^j = (x_1^j, x_2^j, \ldots, x_T^j)$ combines $T = 50$ frames. And one frame

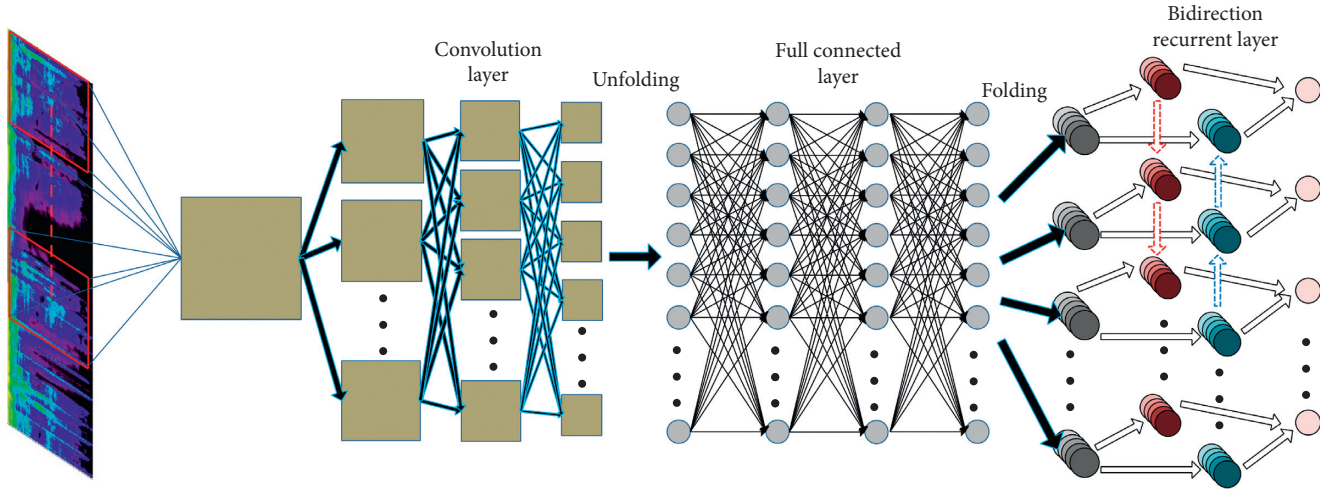


Figure 1: The deep learning framework of our model. The model is a bidirectional RNN, which begins with three convolutional layers (the third layer is mean pooling) and is followed by four fully connected layers. The data are translated into spectrogram as input to the network and are trained end to end.

$x_t^j = (x_{t,1}^j, x_{t,2}^j, \ldots, x_{t,f}^j)$ is a $f = 1764$ dimension vector $(44100/(1000/40) = 1764)$. Thus, a voice recording of 136 seconds can be divided into $n = 68$ samples $(136/2 = 68$, 40 ms$*50 = 2$ s). Each sample x^j has a corresponding label sequence $y^j = (y_1^j, y_2^j, \ldots, y_T^j)$, $y_t^j = \{0, 1\}$, in which a frame without breathing is set to 0, and otherwise is set to 1. Generally, the goal of our processing is to convert an input sequence x^j into a 0-1 sequence.

The data described above are suitable as input for a RNN. However, our model is a RNN with several convolutional input layers, which requires the input to be in a two-dimensional structure. Thus, we split one 40 ms frame into 4 frames with 10 ms. Each 10 ms frame is translated from time domain to frequency domain through FFT, which produces a 220-dimension vector. Now, we translate a one-dimensional 40 ms frame into a two-dimensional spectrogram with the size of 220$*$4.

The main idea of our scheme is to differ the breathing events from the low SNR recordings. It needs to support the high-frequency signals for learning the fine-grained features in deep learning model. Thus, we use the sampling rate of 44100 Hz, which is the highest sampling rate of most smartphones on the market.

3.2. Batch Normalization for Deep Bidirectional RNNs. To efficiently absorb data, we increase the depth of the network by adding more convolution and fully connected layers. However, it becomes more challenging to train the network using gradient descent as the size and the depth increase; even the Adagrad algorithm could achieve limited improvement. We add batch normalization [22] to train the deeper network faster. Recent research has shown that batch normalization can speed convergence, though not always improving generalization error. In contrast, we find that when applied to very deep RNNs, it not only accelerates training but also substantially reduces final generalization error.

In a typical feed-forward layer containing an affine transformation followed by a nonlinearity $f(\cdot)$, we insert a batch normalization transformation by applying $f(Wh + b) \to f(\mathscr{B}(Wh))$, where

$$\mathscr{B}(x) = \gamma \frac{x - E(x)}{(\mathrm{Var}[x] + \varepsilon)^{1/2}} + \beta, \tag{1}$$

in which the terms E and Var are the empirical mean and variance over a minibatch, respectively. The learnable parameters γ and β allow the layer to scale and shift each hidden unit as desired. The constant ε is small and positive and is included only for numerical stability. In our convolutional layers, the mean and variance are estimated over all the temporal output units for a given convolutional filter on a minibatch. The batch normalization transformation reduces *internal covariate shift* by insulating a given layer from potentially uninteresting changes in the mean and variance of the layers' input.

A recurrent layer is implemented as

$$\begin{aligned} \overrightarrow{h}_t^l &= f\left(W^l h_t^{l-1} + \overrightarrow{U}^l \overrightarrow{h}_{t-1}^l + b^l\right) \\ \overleftarrow{h}_t^l &= f\left(W^l h_t^{l-1} + \overleftarrow{U}^l \overleftarrow{h}_{t+1}^l + b^l\right), \end{aligned} \tag{2}$$

where $\overrightarrow{h}_t^l$ and $\overleftarrow{h}_t^l$ are computed sequentially from $t = 1$ to $t = T$ and from $t = T$ to $t = 1$, respectively. And the $l + 1$ (nonrecurrent) layer takes both the forward and backward units as inputs $h_t^{l+1} = f(W^{l+1} h_t^l + b^{l+1})$, where $h_t^l = \overrightarrow{h}_t^l + \overleftarrow{h}_t^l$, and the activation function $f(x) = \min\{\max\{0, x\}, 20\}$ is the clipped ReLu.

There are two ways of applying batch normalization to recurrent operation:

$$\begin{aligned} h_t^l &= f\left(\mathscr{B}\left(W^l h_t^{l-1} + U^l h_{t-1}^l\right)\right) \\ h_t^l &= f\left(\mathscr{B}\left(W^l h_t^{l-1}\right) + U^l h_{t-1}^l\right), \end{aligned} \tag{3}$$

where the first one indicates that the mean and variance statistics are accumulated over a single time step of the

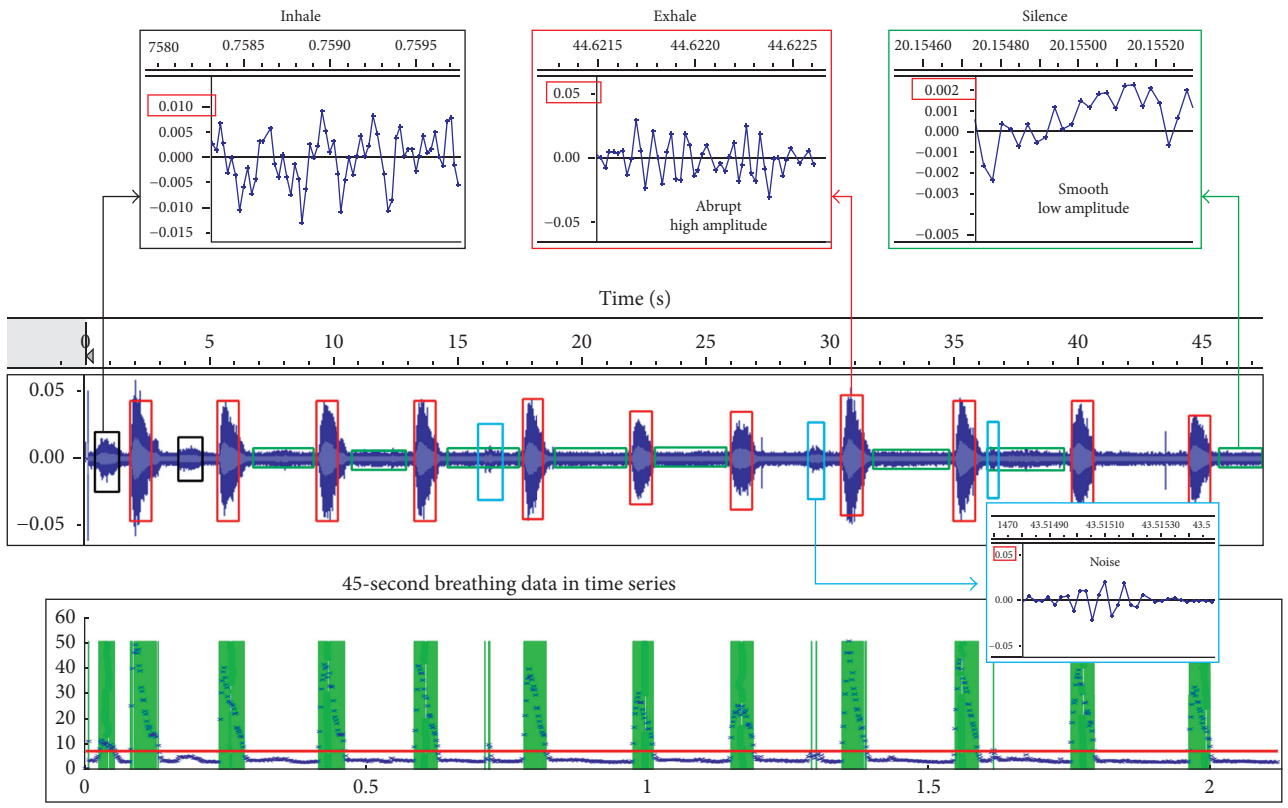


FIGURE 2: A snapshot shows how we label the data semiautomatically, which is motivated from the observation of the frequency and the amplitude of the data.

minibatch, which is ineffective in our study. We find that the second one works well, and Cooijmans et al. [23] have explained the reason.

3.3. Convolutions. Temporal convolution is commonly used in speech recognition to efficiently model temporal translation invariance for variable length utterances. Convolution in frequency attempts to model spectral variance due to speaker variability more concisely than what is possible with large fully connected networks. We experiment with the use of convolutional layers from one to three. These are both in the time and frequency domains (two-dimensional) and in the time-only domain (one-dimensional).

Some previous works [24] demonstrated that multiple layers of two-dimensional convolution improve results substantially than one-dimensional convolution. And convolution has good performance on noisy data. A key point and necessity of low SNR data recognition is denoise [24]. Thus, the convolution component of the model is the key point for it to work well on low SNR data.

3.4. Bidirectional. Recurrent model with only forward recurrences routinely performs worse than similar bidirectional ones, so that implying some amount of future context is vital for good performance. However, bidirectional recurrent models are challenging to be deployed in an online, low-latency setting because they cannot stream the transcription process as the utterance arrives from the user. The DeepSpeech2 of Baidu [24] supports a special layer called lookahead convolution, which is a unidirectional model but without any loss in accuracy comparing to a bidirection one. In our study, we use the bidirectional recurrent model because we do not have the constraint of working in real time.

4. Fine-Grained Breathing Rate Monitoring

4.1. Training Data. Large-scale deep learning systems require an abundance of labelled data. For our system, we need a lot of labelled low SNR data, but it is difficult to label the low SNR data. To train our large model, we first collect high SNR data (the data collected in a quiet surrounding in which breathing can be heard clearly), which are much easier to be labelled. After that, we combine the labelled data with pink noise to lower the SNR of the data through data synthesis, which has been successfully applied to data extensions of speech recognition.

Figure 2 is one high SNR sound recording of 45 seconds. The 11 red boxes mark 11 expirations, and the black, green, and blue boxes mark 2 inspirations, 8 silences, and 3 noises, respectively. When we enlarge the frames of expiration, inspiration, silence, and noise, it is obvious that the differences between adjacent points in a silence frame is much lower than that in the other three types of frame on average.

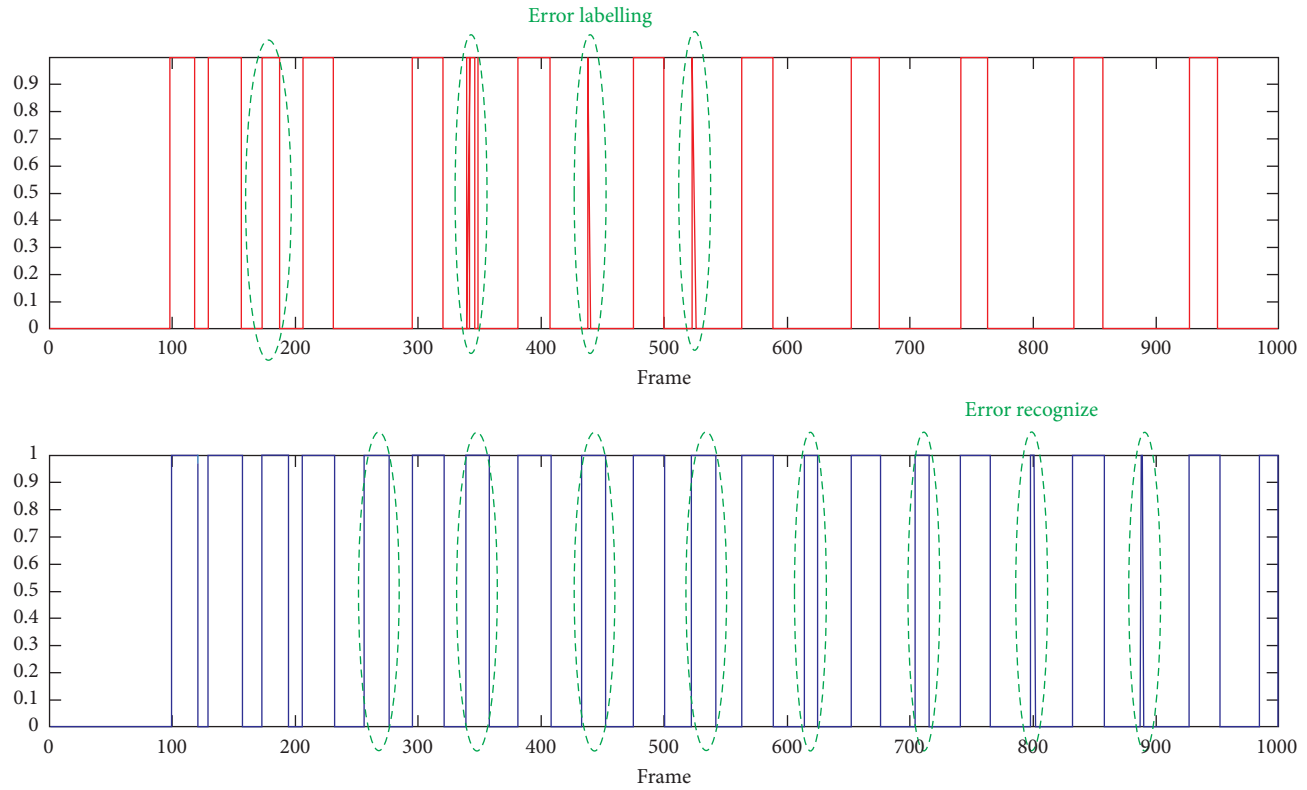


Figure 3: Postprocessing. A comparison between truth and test which shows the necessity of postprocessing.

And its amplitudes are also minimum, only 10^{-3}. Then, we give each frame a frame index $l(x^i)$ as follows:

$$l\left(x^i\right) = \sum_{t=1}^{n} \left|x_t^i\right| + \left|x_{t+1}^i - x_t^i\right|, \tag{4}$$

where x^i is the ith frame in data, $n = 1764$. The subfigure on the bottom of Figure 2 shows the label indexes of one sound recording (x-axis is frame and y-axis is the corresponding frame index). To facilitate observation, we add green lines on the frames whose index is larger than the threshold. It makes the label indexes to distinguish the silence frame from the other three types of frame well. For labelling the data, we erased the inspiration and noise manually (since it is a high SNR data, the volunteers are required to keep quiet during recording), and the threshold is also given manually. All the above actions are conducted in a sound processing software *Audacity* [25], and the thresholds and labels are determined on *Matlab*.

It took us 3 weeks for labelling about 7458 sound recordings (one recording is about 2 minutes, and the total recordings from 109 people are about 248 hours). Finally, we use *Audacity* to complete the data synthesis discussed above.

During the labelling, we find that the number of breathing frames is larger than that of nonbreathing frames, which may reduce the classification performance. Thus, we add a weight in loss function as follows:

$$Ł\left(\widehat{y}, y\right) = \frac{1}{T} \sum_{t=1}^{T} \left[\eta y^t \ln \widehat{y}^t + \left(1 - y^t\right) \ln\left(1 - \widehat{y}^t\right)\right]. \tag{5}$$

4.2. Postprocessing. Figure 3 is a snapshot of 1000 continuous frames randomly chosen from training data, in which the top figure is the ground truth, while the bottom figure shows the recognition results from our model (the x-axis is the frame and the y-axis is the label value). As shown in the figure, breathing is continual and periodical, while the incorrectly labelled data and false-recognized frames are abrupt and discrete. Thus, we can define postprocessing as follows:

First, we can regard a breathing event as the continuous frames with breathing label. Thus, we delete the breathing events whose frame number is less than a threshold, which is the key point of postprocessing. In this study, we choose a value of 50 because a breathing time less than 0.5 s is abnormal according to our tests. As shown in Figure 3, the postprocessing could remove the green dotted-line cycle. The effect of postprocessing is shown in Table 2,and the values of TTR (Test ground-Truth Rate) can demonstrate its efficiency (TTR is a metric described in Section 5.3).

5. Experiment

5.1. Training. We train our deep RNN on 2 work stations with the configuration as follows: Intel Xeon processor E5-2620, 32 GB memory, standard AnyFabric 4 GB ethernet adapter, and Nvidia GTX Titan GPU with 12 GB memory. We use a PC as *parameter server* and 2 work stations as *workers*. Each individual machine sends gradient updates to the centralized parameter repository, which coordinates

Table 1: The results from 6 models: support vector machine (SVM), logistic regression (LR), DRNN (baseline), Model 1, Model 2, and Model 3. All of them are trained and tested on two data sets. The rightmost 4 columns are the average scores over the 4 distances.

	20 cm (0.8 m)		40 cm (0.95 m)		60 cm (1.1 m)		80 cm (1.25 m)					
	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	WAR	TTR (bmp)
SVM	0.897	0.923	0.787	0.920	0.402	0.843	0.316	0.820	0.601	0.876	0.807	0.53
LR	0.867	0.913	0.757	0.932	0.413	0.789	0.342	0.829	0.594	0.865	0.798	0.53
Baseline	0.994	**0.988**	0.948	0.916	0.509	0.925	0.461	0.947	0.715	0.944	0.886	0.64
DeepFilter 1	0.996	0.970	0.980	0.986	0.667	**0.932**	0.639	0.921	0.820	**0.952**	0.919	0.71
DeepFilter 2	**0.998**	0.965	0.981	**0.988**	**0.740**	0.910	0.703	0.900	0.855	0.943	0.921	**0.77**
DeepFilter 3	**0.998**	0.976	**0.998**	0.976	0.733	0.900	**0.710**	**0.921**	**0.859**	0.943	**0.922**	**0.77**

these updates and sends back updated parameters to the individual machines running the model training. We use a public deep learning library Tensorflow[?] to implement our system.

There are four deep learning models trained in our study. The baseline model is a unidirectional RNN with 4 hidden layers, and the last hidden layer is a recurrent layer. The framework is 882-2000-1000-500-100-1, and learning rates $\alpha = 10^{-5}$ without momentum. Training data are described in the previous section, and we take 40 ms as a frame and 50 frames as a group ($T = 50$). Other three models are also RNNs but begin with convolution layers. The detailed model parameters are listed in Table 2. The third line in the table is the number of hidden layers, respectively. The convolution in one dimension and that in two dimensions has different inputs. For the one-dimensional convolution, the input is a 40 ms frame and is translated to a frequency domain with 882 dimensions. The input of the two-dimensional convolution is a 40 ms frame too and is translated into a spectrogram (4×220). Then, one or two convolutional layers are followed with a mean pooling layer, and the mean pooling size is 3×2 (4×4 for one-dimensional mean pooling). All models have three fully connected layers, and each layer has 512 units. They are ended with two unidirectional recurrent layers except DeepFilter 3, which is ended with one bidirectional recurrent layer. All the models are trained through Adagrad with an initial learning rate $\alpha = 1e^{-3}$.

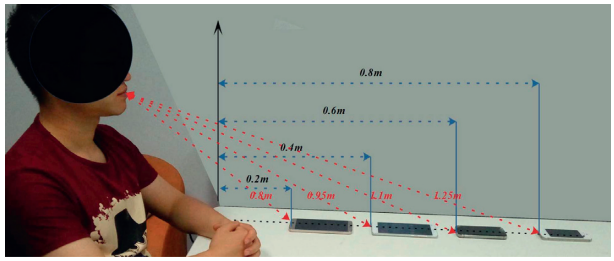
Table 2: Model parameters. The convolutional layer parameters are denoted as "⟨convolutional type⟩ conv ⟨receptive field size⟩ – ⟨number of channels⟩." Also, the parameters of recurrent layer are denoted as "⟨uni or bi⟩ – ⟨number of hidden units⟩," where "uni" denoted unidirectional recurrent and "bi" denoted bidirectional recurrent. The ReLU activation function is not shown for brevity.

Model configuration		
DeepFilter 1	DeepFilter 2	DeepFilter 3
6 weight Layers	7 weight Layers	6 weight Layers
Input (882×50 sequence)	Input ($220 \times 4 \times 50$ spectrogram)	
One-dimension-conv9-16	Two-dimension-conv3-16	Two-dimension-conv3-16
	Two-dimension-conv3-32	Two-dimension-conv3-32
	Mean pool	
	FC-512	
	FC-512	
	FC-512	
Uni-128	Uni-128	Bi-128
Uni-128	Uni-128	—
	$\alpha = 1e^{-3}$	
	Output: sigmod	
	Loss: cross-entropy	

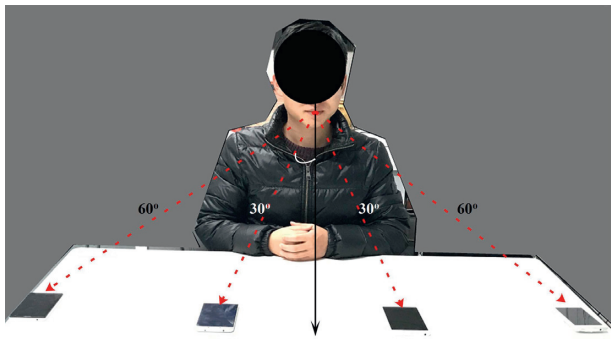
5.2. Experimental Data. Figure 4 shows the procedure of data collection. In Figure 4(a), the volunteer sits in front of the desk, and four smartphones are placed on the desk with a distance of 0.2 m, 0.4 m, 0.6 m, and 0.8 m from the margin of the desk, respectively (the distances from the volunteer's nose to smartphones are far enough, which reach 0.6 m, 0.85 m, 1.0 m, and 1.2 m, resp.). The further the distance is, the lower the SNR is, and the more difficult the labelling of data is. We make four smartphones to possess the same label by collecting data in synchronization, while it is easy to label the nearest one. In Figure 4(b), the volunteer sits in front of the desk, with four smartphones on the desk with a distance of 0.4 m from the margin of the desk, with 4 different angles ±30° and ±60°. We collect 10 volunteers' breathing data, each including 2 minutes tests as Figures 4(a) and 4(b), respectively, and label them finally.

We find some differences on the smartphone with different manufacturers. A funny discovery is from iPhone 4s. iPhone 4s has much worse ability to collect sound recording like breathing, since the built-in microphone has a filter function that can filter the low-frequency noise before recording. This function is developed to improve the quality of voice conversations. We test some smartphones from different manufacturers such as OnePlus, Huawei Honor, Mi, Meizu, Oppo, and Vivo and finally find that 4 Mi smartphones can collect more intact data. Consequently, we choose 4 Mi smartphones with the same band, for removing the unexpecting hardware diversities in experiments.

5.3. Results. There are four metrics to measure the performance of an algorithm in Table 1. TPR (true positive rate),



(a)



(b)

Figure 4: Collecting the test data. (a) Four smartphones in different distances. (b) Four smartphones in different angles.

TNR (true negative rate), WAR (whole accuracy rate), and TTR. There are two classes of samples in our data set: breathing frames (positive samples) and nonbreathing frames (negative samples). TPR is a recognition accuracy on positive samples, while TNR is a recognition accuracy on negative samples. And WAR is the recognition accuracy on the whole data set. After recognizing from deep learning model, breathing frames are calculated into TTR. Here, TTR is a measure of breathing rate, which is defined as a/b, where "a" is the breathing rate calculated by postprocessing and "b" is the ground-truth breathing rate. Table 1 also lists the TPR and TNR of four distances respectively, and it shows that the number of positive samples is quadruple the number of negative ones. That is to say the recognition accuracy of negative samples is much higher than that of positive samples.

We can obtain five results from Table 1 as follows: first, the deep learning models exhibit advantages on precise recognition in comparison to SVM and LR. The superiority increases with the decrease of SNR of the data. Second, the convolution exhibits good performance according to the results of our models and baseline. The two-dimensional convolution is better than the one-dimensional convolution, which demonstrates the ability of feature representation of convolution. And, the frequency domain (two-dimensional) provides more information than the time domain (one-dimensional). It is said that the convolutional layer brings the most improvement on accuracy of recognition. Third, the bidirectional recurrent layer is better than the unidirectional one, according to the results of DeepFilter 2 and DeepFilter 3. It means that not only the history information but also the future information can improve the accuracy of recognition. In practice, the improvement of the bidirectional recurrent layer is limited, which is not much than that of convolution. Fourth, the results of baseline, DeepFilter 1, DeepFilter 2, and DeepFilter 3 demonstrate that the convolution is performed well in both time and frequency domains. And the performance of negative samples (TNR) is much better than that of positive samples in deep learning models. Finally, DeepFilter 3 obtains the best result in most cases, especially for lower SNR data. And the results from DeepFilter 1 to DeepFilter 3 demonstrate that the convolutional layer and the recurrent layer are not conflicts but boost each other in our problem.

In Table 1, there are significant differences between 40 cm and 60 cm with respect to the recognition accuracy. Since the SNR of breathing recordings decreases exponentially with the increasing of distance, TTR can directly indicate the performance of fine-grained breathing rate monitoring. As we see, the TTR of DeepFilter 3 is close to that of DeepFilter 2, and the TTR values of 6 algorithms are less than other three metric values. It implies that one breathing event is separated by some misclassified frames, in which each part is less than 5 frames. Since such breathing events are removed by postprocessing, the TTR values of 6 algorithms are less than other three metric values.

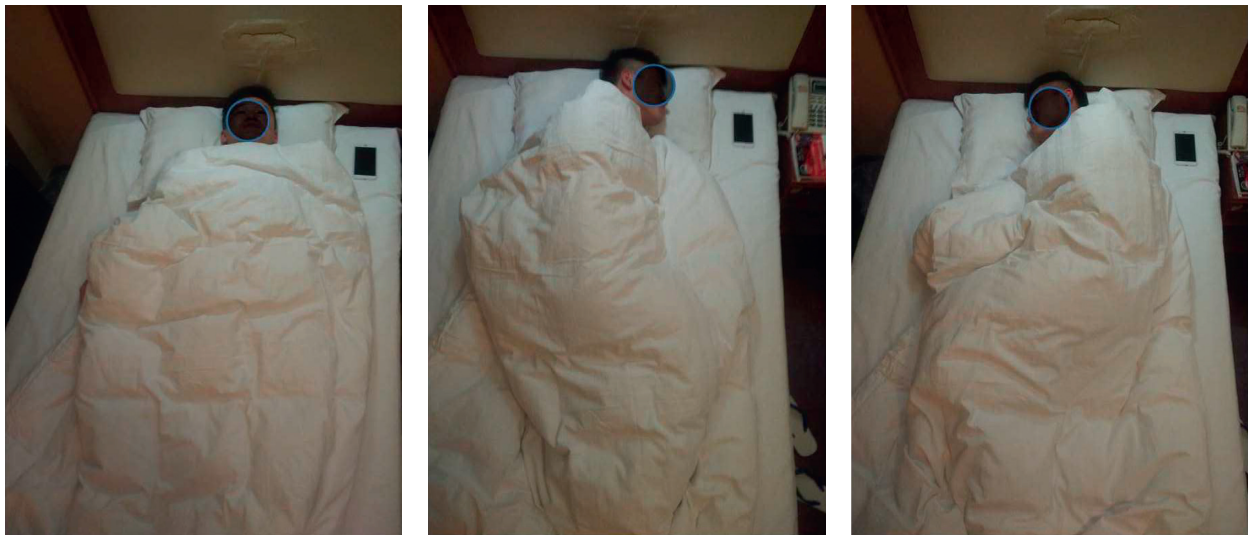
Table 3 lists four related vital sign monitoring methods, including ApneaAPP [17], FMBS [10], WiFi-RF [11], and DoppleSleep [13]. All the four vital sign monitoring methods belong to contactless sleep monitoring systems, which involve breathing frequency detection. The column "Modalities" lists the method of the system used. "RiP-L" and "RiP-U" are the lower bound and the upper bound of accuracy of the system in the corresponding papers, and "bmp" is the unit which means the difference between the rate detected by the system and the actual rate. "RiO-TTR" is the result that we reproduce in the four methods with our data on the metric "TTR." Different from our work, WiFi-RF and DoppleSleep need extra device to assist breathing monitoring while our system only requires the off-the-shelf smartphone. ApneaAPP uses frequency-modulated continuous wave (FMCW) to capture the breathing frequency. We reproduced the method but obtained poor results (as shown in the first line of the last column of Table 3, < 0.2 means less than 0.2). The reason lies in that FMCW needs the device to transmit ultrasonic, and the higher the frequency of ultrasonic is used, the better the results are found. Most smartphones only support ultrasonic less than 20 kHz, which cannot provide enough accuracy in practical (only 0.2 accuracy can be achieved in our reproduced scheme, which is much lower than the 0.98 listed in paper). FMBS is the only method using voice recording, which is most similar to our work. FMBS uses earphone to reinforce voice recordings of the users during the night and adopts envelope detection to assist the breathing detection. But, it is a coarse-grained breathing detector, which only achieves a low TTR value as 0.53 when it is running on our voice recordings. As so far, there are no any other works using smartphone on fine-grained breathing rate monitoring of voice recordings. WiFi-RF and DoppleSleep use WiFi radio frequency and extra device to capture breathing, respectively, which may

Table 3: The comparisons of several breathing rate monitoring techniques. "RiP-L" and "RiP-U" are the lower and upper bounds of accuracy in the papers, respectively. "RiO-TTR" is the result that we obtained by reproducing the methods with our data on metric "TTR."

Related methods	Modalities	RiP-L (bmp)	RiP-U (bmp)	RiO-TTR (bmp)
ApneaAPP [17]	Ultrasonic	<0.98	>**0.996**	>0.2
FMBS [10]	Voice recording	<0.5	>0.95	0.53
WiFi-RF [11]	Radio frequency	<0.6	>0.8	—
DoppleSleep [13]	Radar module	<0.786	>0.893	—

Table 4: The results of four angles.

Angles	±30°	−30°	±60°	−60°
Accuracy ratio	95.01%	94.50%	91.67%	92.81%



(a) (b) (c)

Figure 5: Sleep test.

achieve acceptable accuracy but are not suitable for breathing monitoring based on smartphones.

We can find that the TTR of DeepFilter is 0.77 in Table 1, which is better than 0.53 of FMBS [10]. It means our deep learning algorithm can achieve comparable performance of several professional devices.

We also list the recognition accuracies of four distances in Figure 4(a) and four angles in Figure 4(b). As shown in Table 1, the accuracy is reducing with the raising of distance. An interesting phenomenon of accuracy is dramatically declined on 80 cm, which may be an explanation of critical point of SNR. We will conduct further study on it in our future work. And the results of accuracy ratio affected by different angles are listed in Table 4, which shows a tiny influence on accuracy by different angles.

The precision of polysomnography requires over 99% which is the standard in sleep quality monitoring. In practice, DeepFilter can achieve 90% accuracy in 1 meter but lower than 80% in 2 meters. The 2-meter monitoring distance is sufficient for most fine-grained breathing rate monitoring applications. In most cases, the apps can work well within 1 meter as monitoring distance.

Table 5: The results of sleep test.

	TPR	TNR	WAR	TTR
Flat	80.00%	92.10%	90.10%	0.77
Right	94.60%	98.80%	97.20%	0.84
Left	78.70%	87.90%	86.20%	0.69

5.4. Realistic Test. Figure 5 shows the procedure of realistic sleep test, and the results are shown in Table 5. The volunteer lies on the bed, and the smartphone is placed at the side of the pillow. The smartphone records the breathing of the volunteer, and another camera records the procedure. The test lasts for 7 hours, and finally we collect enough dirty data. It includes the snore, the voice of turning, and other unknown voices from outside. We choose three voice clips that include three poses shown in Figure 5 and obtain the relatively clean data (the environment is usually quite during sleep, so that high SNR data are easy to find). The three clips last for 23, 25, and 22 minutes, respectively. We labelled the data by hand and run our model to validate the effectiveness. The results shown in Table 5 prove that our method achieves

fine-grained breathing rate monitoring in realistic sleep. And we will conduct more realistic experiments on smartphones and other mobile devices in the near future.

6. Conclusion

In this paper, we try to apply deep learning as fine-grained breathing rate monitoring technique to smartphones for people's daily sleep quality monitoring. We propose DeepFilter, a bidirectional RNN stacked with convolutional layers and speeded up by batch normalization, to perform this task. The desirable results of experiments and realistic sleep test prove that our model achieves professional-level accuracy, and deep learning-based breathing rate monitoring apps on smartphones are promising. Our work also extends the use of deep learning to low SNR data recognition, which is inspiring for more data-processing applications.

In our future work, we will exploit more data to train DeepFilter. It implies that DeepFilter needs to suit more smartphones from different manufacturers. And more robust algorithms of postprocessing should be developed. Then, we will also try more deep learning models to solve our problem. And we will deploy the approaches on smartphones and other mobile devices in the near future.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] World Health Organization, "mHealth: new horizons for health through mobile technologies," *Social Indicators Research*, vol. 64, no. 7, pp. 471–493, 2011.

[2] T. Arnold, "Google glass," *Alaska Business Monthly*, vol. 56, pp. 1307–1321, 2013.

[3] U. V. Albrecht, U. Von Jan, J. Kuebler et al., "Google glass for documentation of medical findings: evaluation in forensic medicine," *Journal of Medical Internet Research*, vol. 16, no. 2, p. e53, 2014.

[4] M. Tomlinson, M. Jane Rotheramborus, L. Swartz, and A. C. Tsai, "Scaling up mhealth: where is the evidence?," *PLoS Medicine*, vol. 10, no. 2, p. e1001382, 2012.

[5] J. M. Parish, "Sleep-related problems in common medical conditions," *Chest*, vol. 135, no. 2, pp. 563–572, 2009.

[6] J. E. Shaw, N. M. Punjabi, J. P. Wilding, K. G. M. M. Alberti, and P. Z. Zimmet, "Sleep-disordered breathing and type 2 diabetes: a report from the international diabetes federation taskforce on epidemiology and prevention," *Diabetes Research and Clinical Practice*, vol. 81, no. 1, pp. 2–12, 2008.

[7] P. X. Braun, C. F. Gmachl, and R. A. Dweik, "Bridging the collaborative gap: realizing the clinical potential of breath analysis for disease diagnosis and monitoring-ctutorial," *IEEE Sensors Journal*, vol. 12, no. 11, pp. 3258–3270, 2012.

[8] G. S. Chung, B. H. Choi, K. K. Kim et al., "REM sleep classification with respiration rates," in *Proceedings of International Special Topic Conference on Information Technology Applications in Biomedicine*, pp. 194–197, Tokyo, Japan, November 2007.

[9] H. M. S. Hossain, N. Roy, and A. Al Hafiz Khan, "Sleep well: a sound sleep monitoring framework for community scaling," in *Proceedings of IEEE International Conference on Mobile Data Management*, pp. 44–53, Pittsburgh, PA, USA, 2015.

[10] Y. Ren, C. Wang, J. Yang, and Y. Chen, "Fine-grained sleep monitoring: hearing your breathing with smartphones," in *Proceedings of INFOCOM 2015*, pp. 1194–1202, Kowloon, Hong Kong, April 2015.

[11] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf wifi," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 15)*, pp. 267–276, New York, NY, USA, 2015.

[12] T. Morgenthaler, C. A. Alessi, D. Bailey, J. Coleman Jr., C. A. Kushida, and M. R. Littner, "Practice parameters for the indications for polysomnography and related procedures: an update for 2005," *Sleep*, vol. 28, no. 4, pp. 499–521, 2005.

[13] T. Rahman, A. T. Adams, R. Vinisha Ravichandran et al., "DoppleSleep: a contactless unobtrusive sleep sensing system using short-range doppler radar," in *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 39–50, Osaka, Japan, September 2015.

[14] T. Hao, G. Xing, and G. Zhou, "iSleep: unobtrusive sleep quality monitoring using smartphones," in *Proceedings of ACM Conference on Embedded Networked Sensor Systems*, pp. 1–14, Rome, Italy, November 2013.

[15] Z. Chen, M. Lin, F. Chen, and N. D. Lane, "Unobtrusive sleep monitoring using smartphones," in *Proceedings of the ICTs for improving Patients Rehabilitation Research Techniques*, pp. 145–152, Brussels, Belgium, 2013.

[16] AppCrawir, "Respiratory rate," 2014, http://appcrawlr.com/ios/respiratory-rate.

[17] R. Nandakumar, S. Gollakota, and N. Watson, "Contactless sleep apnea detection on smartphones," in *Proceedings of International Conference on Mobile Systems, Applications, and Services*, pp. 45–57, Florence, Italy, May 2015.

[18] A. R. Mohamed, G. E. Dahl, and G. Hinton, "Acoustic modeling using deep belief networks," *IEEE Transactions on Audio Speech and Language Processing*, vol. 20, no. 1, pp. 14–22, 2012.

[19] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645–6649, Vancouver, Canada, May 2013.

[20] A. Hannun, C. Case, J. Casper et al., "Deep speech: Scaling up end-to-end speech recognition," *Computer Science*, 2014, in press.

[21] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, long short-term memory, fully connected deep neural networks," in *Proceedings of IEEE International Conference on Acoustics*, pp. 4580–4584, Queensland, Australia, 2015.

[22] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *International Conference on Machine Learning*, Lille, France, July 2015.

[23] T. Cooijmans, N. Ballas, Laurent, and A. Courville, "Recurrent batch normalization," 2016.

[24] D. Amodei, R. Anubhai, E. Battenberg et al., "Deep speech 2: end-to-end speech recognition in english and mandarin," in *International Conference on Machine Learning*, New York, NY, USA, June 2016.

[25] Audacity, "Audacity," 2015, http://www.audacityteam.org/.

Hindawi
Mobile Information Systems
Volume 2018, Article ID 8198379, 12 pages
https://doi.org/10.1155/2018/8198379

Research Article

Improving Learning Tasks for Mentally Handicapped People Using AmI Environments Based on Cyber-Physical Systems

Diego Martín,[1] **Borja Bordel**,[1] **Ramón Alcarria**,[2] **and Yone Castro**[3]

[1]*ETSI Telecommunications, Universidad Politécnica de Madrid, Av. Complutense 30, 28040 Madrid, Spain*
[2]*ETSI Topography, Geodetics and Cartography, Universidad Politécnica de Madrid, Camino de la Arboleda s/n, 28031 Madrid, Spain*
[3]*Fundación PRODIS, Bulevar Indalecio Prieto 2, 28032 Madrid, Spain*

Correspondence should be addressed to Diego Martín; diego.martin.de.andres@upm.es

Received 28 July 2017; Accepted 26 November 2017; Published 28 January 2018

Academic Editor: Pino Caballero-Gil

A prototype to improve learning tasks for mentally handicapped people is shown in this research paper using ambient intelligence techniques and based on cyber-physical systems. The whole system is composed of a worktable, a cyber-glove (both with several RFID and NFC detection zones), and an AmI software application for modeling and workflow guidance. A case study was carried out by the authors where sixteen mentally handicapped people and 3 trainers were involved in the experiment. The experiment consisted in the execution of several memorization tasks of movements of objects using the approach presented in this paper. The results obtained were very interesting, indicating that this kind of solutions are feasible and allow the learning of complex tasks to some types of mentally handicapped people. In addition, at the end of the paper are presented some lessons learned after performing the experimentation.

1. Introduction

Cyber-physical systems (CPS) are intersections of the physical and the cyber worlds [1]. Although different definitions and proposals may be found [2], basically, these new systems allow creating feedback control loops where sensors and actuators obtain information about the real world, which is processed by new generation cyber components (as, e.g., enhanced pattern-recognition techniques), in order to infer the state of the environment and its inhabitants and be able to act to reach a more sustainable and profitable situation [2].

With these ideas, one of the most common applications for CPS is traceability [3]. Both product traceability and people surveillance have been enhanced through the use of CPS in the last years [4]. Moreover, some pedagogical works have proved that this kind of systems are useful to control the student learning in the so-called Industry 4.0 educational scenarios. However, these solutions require a high number of deployments, a complete remodelation of teaching methodologies, and manners. Therefore, usually they are referred as "the systems of the future," discarding their viability nowadays.

Nevertheless, some small deployments or interventions may be useful in certain scenarios, such as in educational environments involving people who are mentally handicapped. These solutions do not fit totally the requirements of the new generation systems but follow their principles employing the existing technology. In the context of ambient intelligence (AmI) systems, for example, this approach is being investigated in a very successful way [5]. In particular, we argue that it is a very promising view for scenarios where traditional techniques are insufficient, but where total automation (as CPS paradigm proposes) is not valid at all (i.e., the human actor must endure). This is the case, for example, of the applications involving people with special needs.

This paper is an extended version of our previous work [6], which proposed an enhanced ambient intelligence (AmI) system based on cyber-physical systems for tracking the tasks being performed by people who are mentally handicapped during their learning process.

The CPS presented in this research work include a cyber-glove, a worktable, and an AmI software for creating and controlling workflows, whose tasks involve actions about the cyber-glove and the worktable. This system is framed in a research project funded by a grant from Madrid Autonomous Community called SEMOLA. Currently, this research project is in a preliminary phase; so the developments done are also at an exploratory stage. The whole system is called LAoCA, an acronym from "Learning Architecture over CPS and AmI."

The objective of this research work is to validate whether an AmI environment for CPS traceability facilitates learning for mentally handicapped people and improves control over tasks and the acceptance of the system by users. Is not the objective of this paper to validate the characteristics of the system with which the experimentation was carried out, which could be a possible future work.

In order to guide this research work, three research questions were stated:

(i) *RQ1*: Can an AmI environment for CPS traceability facilitate learning for mentally handicapped people?

(ii) *RQ2*: Does it improve control over the performed tasks?

(iii) *RQ3*: Is LAoCA accepted by the final users?

A case study was carried out by the authors in order to validate the environment and the solution proposed in this paper, where sixteen mentally handicapped people with Down syndrome from one of the most important foundations for mentally handicapped people in Spain were involved. The validation consisted of studying the improvement of participants' learning capabilities, analyzing the system's control of the tasks performed using the cyber-glove and the worktable, and performing a survey about the final user acceptance of the system.

The results obtained are very promising confirming that the proposed solution is feasible, but as it is a preliminary version. This first validation allowed us to find weaknesses in the proposal with the aim of improving it, in following versions and improving the design validation for future experiments.

2. Related Work

As said before, cyber-physical systems paradigm fits perfectly the requirements of traceability solutions [3]. In fact, various previous works have investigated the use of CPS (and other similar solutions such as Smart Environments [7]) in AmI solutions and traceability systems.

In particular, the National Institute of Standards and Technology (NIST) has proposed a reference architecture valid to be employed as framework for traceability solutions and industrial automatic systems [8]. This proposal aims to turn into the basic document for the future commercial CPS. For example, following this line, some research works have investigated the use of low-energy devices in order to display information in different scenarios such as supermarkets, production systems, or hospitals [9, 10]. The basic idea is to employ new, open, and more efficient technologies such as the electronic ink in order to improve the current automatic industrial systems based on proprietary technologies [11, 12], programmable logic controller (PLC) [13, 14], and very low-level protocols such as OPC (Object Linking and Embedding for Process Control) [15].

On the contrary, CPS approach refers an integrated solution including all the requirements needed for automated industrial processes, allowing a reduction in the system's complexity, and therefore, reducing the investment. Using this framework, some traceability systems have been proposed [3], proving that the number of errors committed by workers gets reduced in comparison to traditional systems. Moreover, solutions focused on inferring the activities being performed by people in certain scenarios (as, e.g., the daily living) have been reported [16]. In these proposals (usually), the enabling technology is radio frequency identification (RFID), which is employed to create cyber-gloves, bracelets, and other instruments [17, 18]. These designs, nevertheless, are usually very obtrusive, so the execution of tasks and processes is highly modified when considering the use of these tools. For this reason, learning solutions have not included this kind of systems yet (as they may impede the development of some activities). This idea, which is valid in general, is much more important and critical when people who are handicapped are considered. Despite the effort of the NIST (and other standardization organizations) to define the concept of CPS, several different proposals and architectures may be found in the literature [2]. Nowadays, even the appearance of the Industry 4.0 systems [19] in Germany (based on CPS solutions and principles) has turned the scene more complex and heterogeneous.

In fact, the Industry 4.0 revolution aims to modify the entire society. This revolution includes learning and teaching. In this context, different proposals about how the future learning systems (based on Industry 4.0 and CPS paradigms) could be constructed and implemented may be found. Systems based on gamification and virtual reality are the most common. However, other proposals based on defining adaptive systems being able to facilitate the learning of, for example, new tasks in industrial scenarios have been reported [20].

Finally, in the last five years, the CPS paradigms have been also applied to learning systems. In particular, as we said, several works try to infer the Activities of Daily Living [21] performed by users using a sort of RFID-enabled cyber-glove [22] (possibly complemented with additional sensors such as accelerometers [23]), connecting the outputs of these tools to a system which helps people to learn in a faster way (by means of real-time feedback, assisted living, etc.). The applications of these systems are mainly focused on occupational therapy [24] for neurological patients; children control, augmented reality [25], and behavioral therapy [26]; and accessibility (such as the system SignAloud proposed by the MIT [27]).

3. LAoCA (Learning Architecture over CPS and AmI)

In this section, we present a second version of our previous work [6], a development framed within a research project

called SEMOLA (2016–2019), funded by a grant from Madrid Autonomous Community. Among the many specific objectives of the SEMOLA project, we developed a prototype to test the technology to support people with intellectual disabilities in their personal development, taking timely remedial actions and anticipating problems that can be found. Thus, we will generate enough knowledge for the upcoming project phases, such as lessons learned and best practices.

Our *LAoCA* (Learning Architecture over CPS and AmI) prototype is composed of three elements: a cyber-glove, a worktable for tracing the items, and an AmI software application for creating, managing workflows, and also controlling the actions performed with the whole system. All components are connected by means of a Bluetooth network with star topology, acting the AmI application as central element. Bluetooth 3.0 technology allows creating ad hoc networks with a valuable data rate (until 3 Mbps), which perfectly fit with the real-time requirements of *LAoCA*. In the next section, we are reviewing in detail each cited component.

3.1. Worktable. In this section, we present a worktable made of plastic with four detection zones as shown in Figures 1 and 2; Figure 3 shows the tagged elements to be detected by the worktable. The main purpose for which the worktable was designed is to identify objects (tagged with RFID or NFC) placed in the four detection zones and transmit that information wirelessly in real time to the control program. It also may detect the moment when an object previously placed over the worktable is removed. The worktable also includes some actuators in order to interact with the users and inform them about the execution of the planned tasks (Section 3.3).

The electronics of the worktable are composed by the following: four printed coils working as antennas which sense the four detection areas; four RDM8800 RFID and NFC readers (one for each detection area); five Arduinos Nano, four dedicated to manage the four antennas and the last one acting as main microcontroller in the system; a bluetooth module for its communication with the control software; four green LEDs (one for each detection zone) used to illuminate the detection surface, and additional three LED indicators (actuators) for showing its operation and giving some feedback to the user. The surface LEDs are used to describe the task to be performed at an earlier stage of training. Figure 4 shows an inside view of the worktable, while Figure 5 shows the electronic schema of the worktable.

3.2. Cyber-Glove. Figure 6 shows the prototype of the cyber-glove proposed in this research paper. It is made of cloth, and it has a compartment where all the necessary electronics for its operation are stored (Figure 7). This compartment can be positioned below or above the wrist in order to facilitate hand movements.

The main functionality of the cyber-glove is to be a device capable of identifying objects tagged with RFID or NFC wirelessly and having communication capabilities with the software that controls and manages the movements. In summary, the cyber-glove may detect any object touched or taken by users and send the information immediately to the

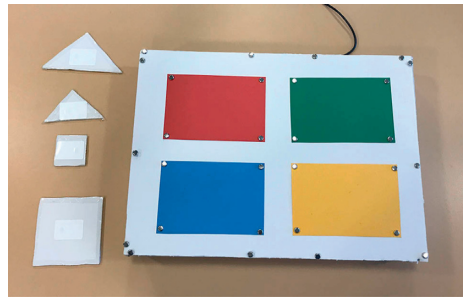


Figure 1: Worktable with four detection zones and NFC tags.

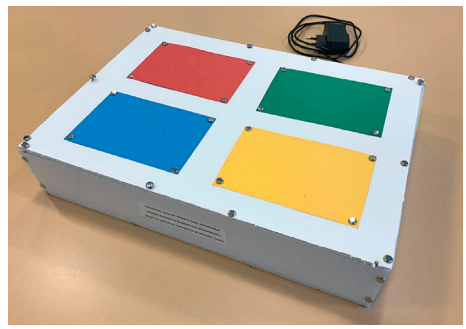


Figure 2: Perspective view of worktable.

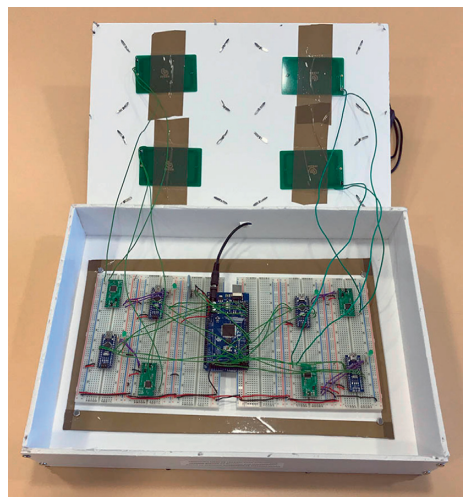


Figure 3: Tagged elements.

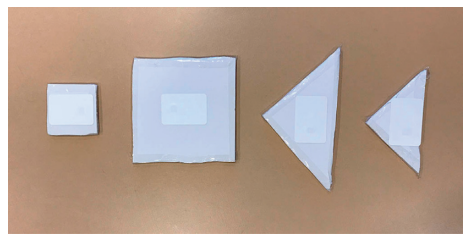


Figure 4: Electronics inside the worktable.

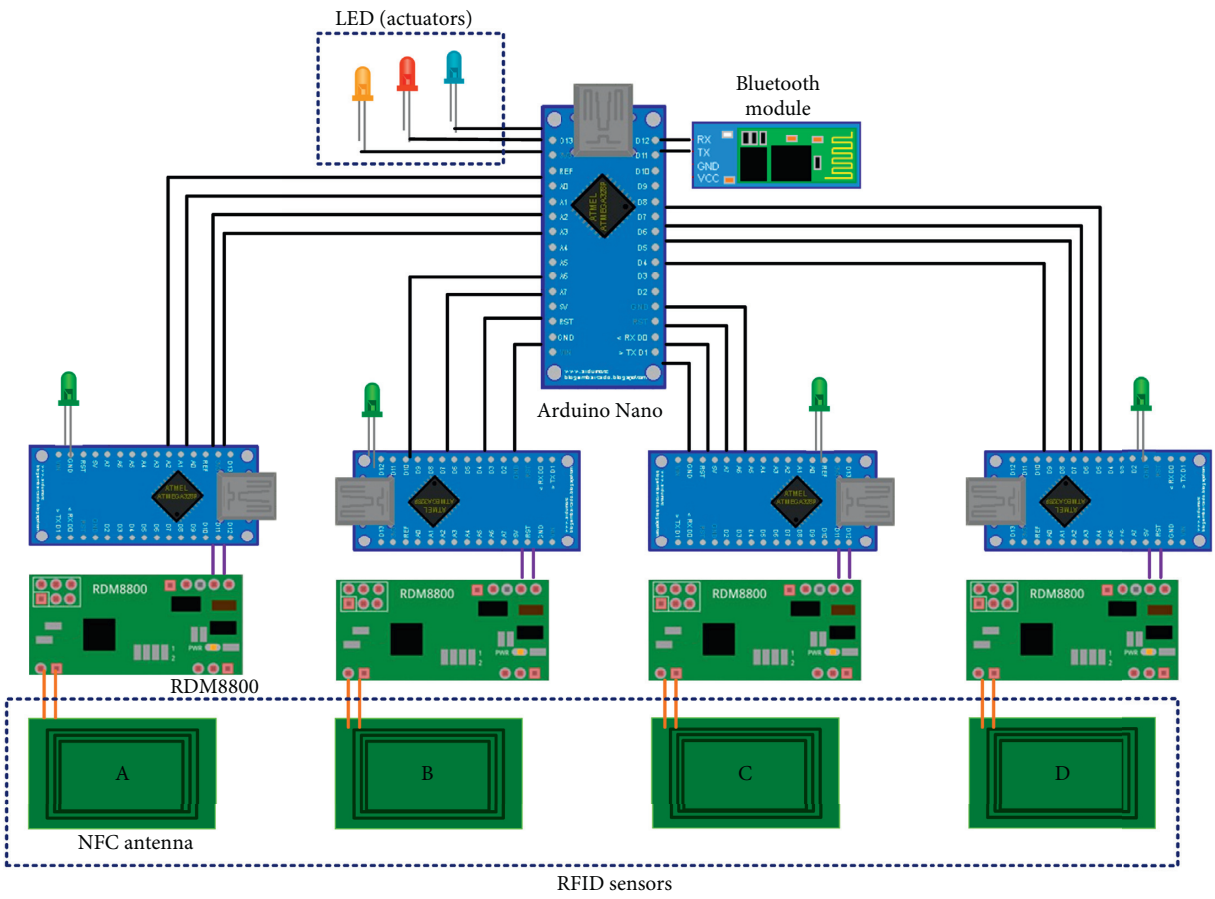


Figure 5: Worktable electronics schema.

Figure 6: Cyber-glove.

management software. The cyber-glove may also detect the moment when the user releases the object (if he held it) informing to the management software. The cyber-glove cannot detect two or more objects touched at the same time.

The electronics of the cyber-glove consists of several elements as seen in Figure 7: an Arduino Nano for data processing, a RDM8800 NFC chip for reading the RFID and NFC elements, an HC-06 Bluetooth module for communicating with the control program, a coil in the palm that works as antenna for reading RFID and NFC elements, and a battery.

3.3. Workflow Creation and Control Software. The workflow creation and control software is an ambient intelligent tool for creating and managing workflows made of different tasks (such as position control or state monitor manage) involving the tagged objects with RFID and NFC through the cyber-glove and the worktable. The different tasks are related by means of some "transitions" which are triggered, if the associated condition is fulfilled.

This software has two main features: (a) it models the workflow of the task intended to execute, using a *prosumer* interface that allows a user with no programming experience to model the tasks, and (b) it also executes the control workflows previously created in order to supervise and analyze the actions performed by a user who is working with the cyber-glove and the worktable.

For workflow creation, the AmI tool includes a graphic environment where users may compose their own workflows using different predesigned modules. In respect to the second functionality (workflow control), the software permanently is hearing for the information from the cyber-glove and/or the worktable. It evaluates the notifications and triggers the proper transitions (if any exists). Every workflow should include "fatal error tasks" to which the workflow moves if the execution fails. Besides, a "successfully finished task" would be advisable to be included. In Figure 8, we present an example workflow, indicating the described desirable structure.

Additionally, the AmI software transmits towards the worktable information about the workflow evolution. Then,

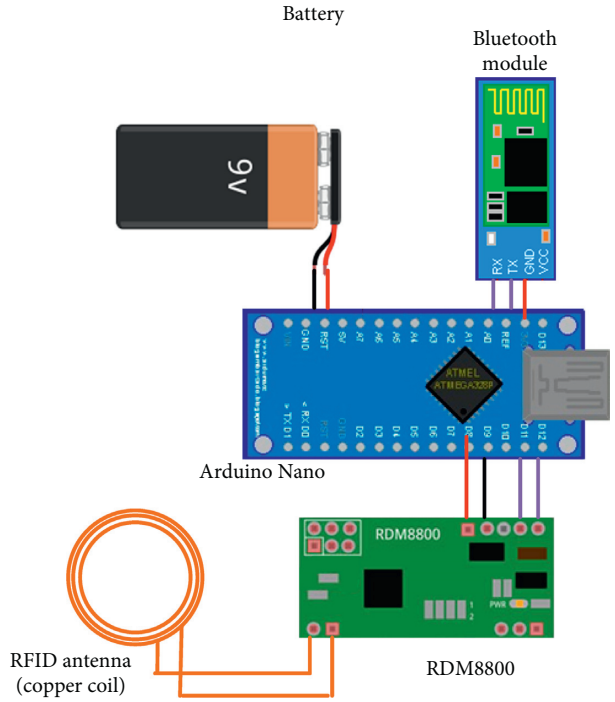


Figure 7: Cyber-glove electronics schema.

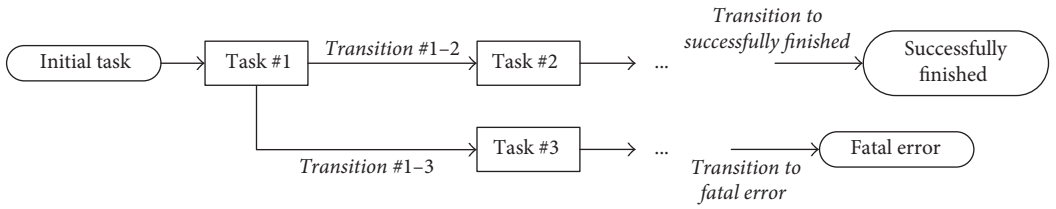


Figure 8: Example workflow.

the worktable can show that information (using the actuators—LEDs) to the users (e.g., if the workflow finishes with a fatal error, a red LED will be turned on).

4. Experimental Validation

This research work has been guided by three research questions presented in the introduction. We designed and performed a case study where sixteen mentally handicapped people performed some tasks, defined and supervised by three trainers from a foundation committed to people with intellectual disabilities. The validation presented in this research work is a preliminary action framed in the early years of development of the SEMOLA project. Our goal is to test the feasibility of LAoCA and, with the results obtained, improve it with a new deployment at the foundation and extend the experimentation with more people involved. The researchers did not influence participants using the environment. This approach is appropriate to replicate the experiment in similar contexts.

4.1. Context. The authors of this research work (hereafter, experts) designed, executed, and assessed the case study, whose research areas are cyber-physical systems, ambient intelligence, knowledge management, etc.

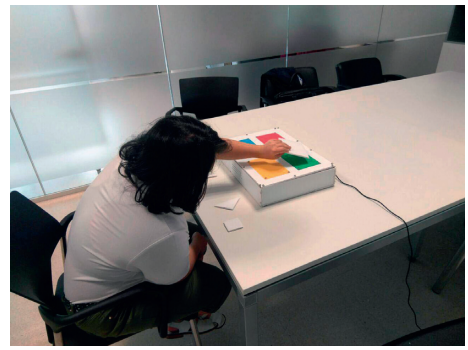
Sixteen mentally handicapped people from the foundation (hereafter, participants) participated in the experimentation performing 4 tasks each, and three trainers from the same foundation helped in training the participants in some tasks using the software for workflow modeling of LAoCA.

Sixteen participants conformed the sample. The sample meets the criteria of accessibility, suitability, and representativeness. All of them are potential users of the system proposed in this research work. The 16 participants were grouped into four groups depending on their type of disability, and the characteristics of the participants are described in Table 1. Participants were selected by the experts to perform the experimentation following the characteristics described in Table 1.

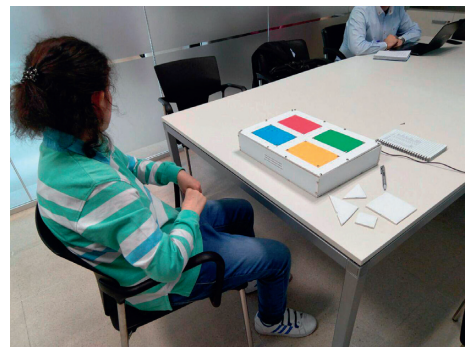
It is considered that the sample is representative since the average age is located in the general middle point of development of people with intellectual disabilities, taking into account that 80% of the sample has Down syndrome. None of the participants developed cognitive impairment (evaluation through CAMCOG [28]). The degree of disability and

Table 1: Summary of the main features of the participants (own elaboration).

Level	Age (mean)	Type of disability	Disability (%) (33–100)	Limitation on the global activity (%) (0–100)	Cognitive impairment (low-moderate-severe)	Oral comprehension level (low-medium-high)	Mobility (difficulty-no difficulty)
1	24.25	Psychic	64	61.75	Moderate-low	Medium high	Difficulty
2	26	Psychic	69.75	64.75	Moderate-severe	High	No difficulty
3	22.75	Psychic sensory	74.25	74.5	Severe-moderate	Medium-low	Some difficulty
4	24.5	Psychic, physical and sensory	87.25	83.25	Severe	Low	Difficulty
Mean	24.38	—	72.75	71.06	—	—	—



(a)



(b)

Figure 9: Two participants performing the experiment.

limitation in the global activity is representative of different stages of dependence, and they reflect what type of difficulties can occur in the general population with intellectual disability susceptible to use this type of technology. Representatives of different levels of oral comprehension have been employed, in order to counterbalance the results in relation to the instructions that have been used in the method. Likewise, people with different types of problems have been chosen (visual and motor) that could affect the test. Figure 9 shows two photographs of two people participating in the experiment.

The test population was selected to be as much heterogeneous as possible in order to validate (in an implicit way) the adaptation capability of the system for different users. In fact, no significant differences were observed between the system performance when employed by different users.

4.2. Planning. The experimentation plan was executed in four phases.

4.2.1. Training Phase. Trainers received some instructions about the use and operation of the CPS presented in the paper with the aim to train the participants about the use of the cyber-glove and to perform some tasks with it. The training sessions consisted of a one-hour lecture at the beginning of the experimental validation given by one of the experts.

Table 2: Statistical summary of the errors committed in the tasks' execution.

Task difficulty level	With LAoCA				Without LAoCA			
	Median	SD	Min	Max	Median	SD	Min	Max
Easy	1	1.26	0	4	3	1.61	1	6
Medium	4	1.01	2	5	4	1.26	2	7
Difficult	5	1.11	3	7	5	1.5	3	8
Total	4	1.76	0	7	4	1.75	1	8

4.2.2. Task Performing. The participants executed six different tasks: three of them were done using LAoCA and the other three without it. At the beginning of each task, the system showed the sequence of movements through a tablet where the participants observed in a visual diagram the movements to be performed with the tagged objects (shown in Figure 3) on the worktable in order to learn the task. For the experiment without LAoCA, the coach was in charge to explain the task to the participants. A task is a sequence of movements over the worktable of several tagged objects. Following the recommendations of the trainers, for this experimentation, we defined three types of tasks depending on its difficulty:

(1) Easy tasks: composed of five movements of only one tagged object
(2) Medium tasks: composed of five movements of each two tagged objects
(3) Difficult tasks: composed of eight movements of each four tagged objects.

All the tasks were randomly generated. The initial idea for experimentation was to use the three elements proposed in this research paper as described in Section 3. But, we had serious problems for the participants to use the glove correctly since most of them paid more attention to the glove than to the accomplishment of the tasks. Some participants completed the tasks with the glove and others did not, which is why we decided to remove the data obtained through the cyber-glove in the experimentation. We understand that the cyber-glove is a very intrusive element for people with mental disabilities; we will discuss this in Section 6 and conclusions.

4.2.3. Data Evaluation Phase. In this phase, the trainers were asked about the participants' impressions about the use of the system by the experts. They also helped to assess 48 tasks performed by the participants.

4.3. Data Collection. The data gathered to answer *RQ1* were obtained from the software that guided the tasks and the notes taken by the experts during the "*Tasks performing phase*." We measured the number of errors made by the participants. For example, a wrong movement or moving the wrong object is considered an error. The data collected to address *RQ2* were obtained from the control software and workflow guidance of LAoCA and the perception from the trainers; at first, a survey was considered to ask the trainers, but due to small number we opted for oral interviews. The information to answer *RQ3* was obtained through surveys filled by the trainers after evaluating each participant.

An experiment was carried out to answer the research questions proposed in Section 1, where different mentally handicapped users, characterized by Table 1, performed different tasks described in Section 4.2.

5. Results

In this section, we present the results obtained from the case study. The data and information will be presented trying to answer the three research questions.

5.1. Can an AmI Environment for CPS Traceability Facilitate Learning for Mentally Handicapped People? This section presents the results from the validation phase where the participants performed 4 tasks: two of them using the proposal presented in this research work and the other two without using our proposal, that is, learning in a traditional way the actions needed to complete the workflow.

Table 2 shows a summary of the data obtained, presenting statistical values on the errors committed in the execution of the tasks. The table presents the median, the standard deviation, the minimum and maximum number of errors for each type of tasks. The median of errors using LAoCA is 4 with a standard deviation of 1.76 and a number of errors ranging from 0 to 7, while without the use of our proposal, the median error is also 4, with a standard deviation of 1.75 and values ranging from 1 to 8.

Figure 10 shows two boxplots comparing the distributions of errors for the tasks guided with LAoCA and without it. As we can see, they have the same value for the medians, and the standard deviation is practically the same for both cases. Nonetheless, we performed a Mann–Whitney U test in order to check whether the participants perform less errors by using LAoCA. The result was negative ($p > 0.05$), indicating that there is no evidence to confirm that with our proposal, fewer mistakes are made with the execution of the tasks, as expected from Figure 10.

Continuing the analysis of the data, we realized that people with a higher level of disability made more errors in general, both using our proposal and not using it. Therefore, we decided to exclude persons with the highest level of disability from the study; that is, we removed from the study the people on level 4 as shown in Table 1. The results are shown in Figure 11. With these new data, the distributions are somewhat different as well as the values for the median.

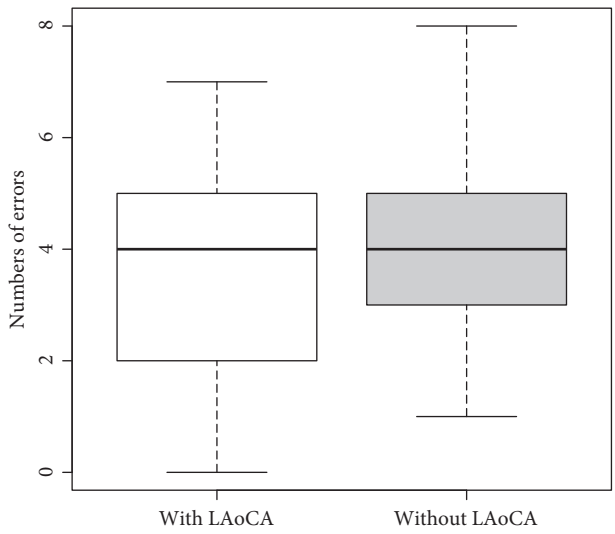


Figure 10: Error distributions for all tasks.

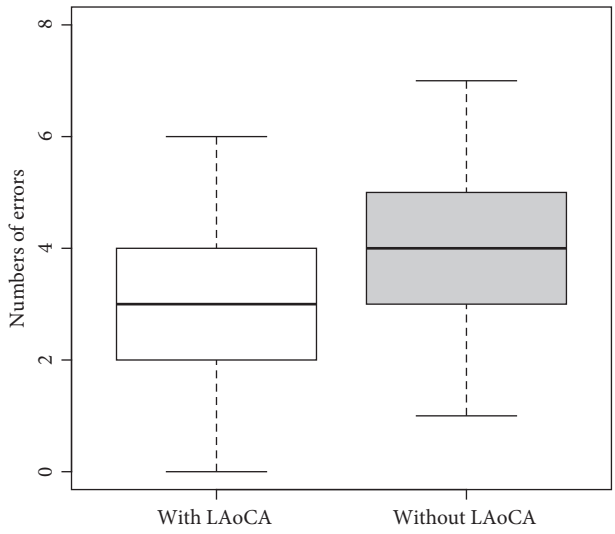


Figure 11: Error distributions for all tasks removing people on level 4.

The median for the group that has used LAoCA is 3 while for the group that has not used LAoCA is 4. We performed a Mann–Whitney U test to confirm if there is now a difference between the distributions, but the results were still negative ($p > 0.05$).

We conducted a more in-depth study, but this time, we decided to exclude people on levels 3 and 4 from the analysis since they were the most vulnerable to making mistakes and keeping the data obtained from people on levels 1 and 2, as shown in Table 1; the data distributions are shown in two boxplots in Figure 12. The medians remain the same as the data distributions shown in Figure 11, 3 for the group that used LAoCA and 4 for those who did not. We performed a Mann–Whitney U test to confirm if there is now a difference between the distributions, and in this case, the result was positive ($p = 0.01$). So, we can state that errors increase when not using LAoCA with people on levels 1 and 2 and all kinds of tasks: easy, medium, and difficult.

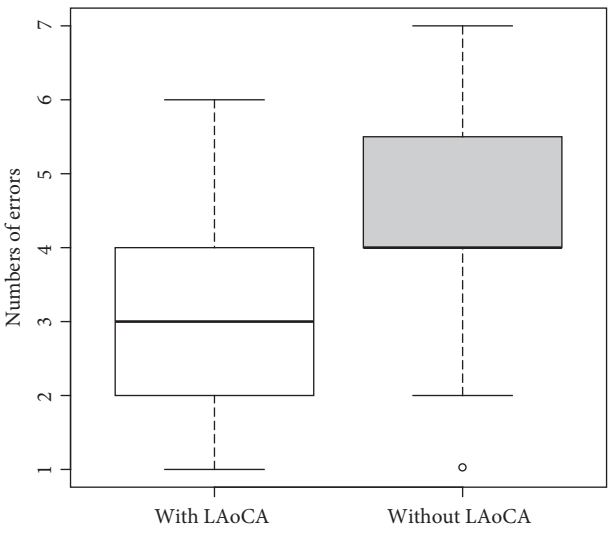


Figure 12: Error distributions for all tasks for the people on levels 1 and 2.

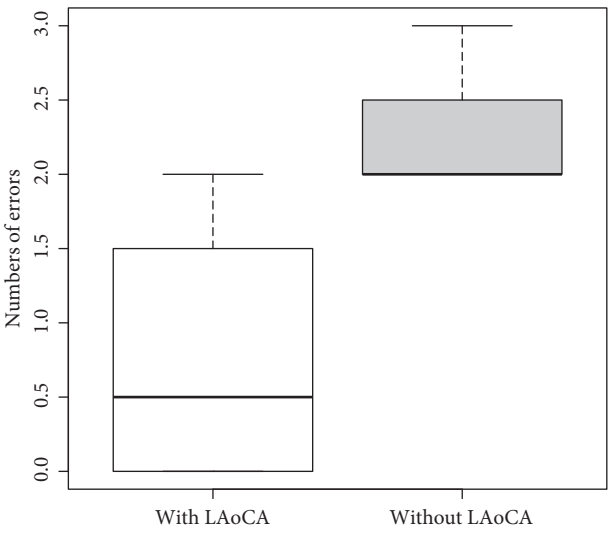


Figure 13: Error distributions for easy task and people on level 3.

However, we performed a further study analyzing the results for each type of task and disability levels. During the execution of the experiment and analyzing the data obtained, we realized that the participants on level 3 also improved their numbers of errors performing the task tagged as easy when using LAoCA; both distributions are shown in Figure 13. In this case, the medians are very different, 0.5 for the case of using LAoCA and 2 for the case of not using it. By performing a Mann–Whitney U test, we confirm the hypothesis that the number of errors was reduced in the case of using LAoCA with the easy task and the people on level 3 ($p < 0.05$). In spite of this, the statistical test performed has very little power because there are very few samples that fulfill the condition of performing the easy task and being in level 3. As will be proposed in the conclusions, this may be a future research work, carrying out a study with more values in the sample.

Table 3: Summary of the tasks improved with the use of LAoCA by levels.

	Easy task	Medium task	Difficult task
Level 1	**	**	**
Level 2	**	**	**
Level 3	*	—	—
Level 4	—	—	—

** $p \leq 0.01$; * $p < 0.05$.

Table 4: Summary of detection errors produced during experimentation.

	Total number of movements	Total number of errors	% of errors
Easy task	80	5	6.25
Medium task	80	7	8.75
Difficult task	128	12	9.375
Total	288	24	8.33

Table 5: Summary of detection errors produced in experimentation excluding persons on level 4.

	Total number of movements	Total number of errors	% of errors
Easy task	60	1	1.67
Medium task	60	2	3.33
Difficult task	96	4	4.17
Total	216	7	3.24

Finally, in Table 3, we show a summary of the tasks that underwent a statistically significant improvement by levels using our proposal.

5.2. Does It Improve Control over the Performed Tasks? In order to answer this research question, we analyzed two types of sources of information. The first one was a study of sensor failures about detecting the movements of tagged objects on the worktable, and the second one was a survey on the opinion of trainers in the execution of the experimentation. Our approach presented in this research work, in very few occasions, had problems to detect the movements of the tagged objects in the worktable. Most of the cases were due to the fact that objects were not placed exactly in the detection zones, and this was due to two main reasons:

(1) The objects are labeled with a small RFID tag in the center of the object, and the detection areas of the worktable have a very specific detection field.

(2) The participants did not place the objects exactly on the detection area.

Since this is a controlled experiment, we were able to measure the detection errors that were produced, below we present a study of these errors. Table 4 shows a summary of the detection errors in all the experimentation presented by tasks; as we can see, the total percentage of detection errors is 8.33, which is above 5% of errors that would be acceptable.

During the experimentation, we detected that people with a higher level of disability produced more errors of detection of the objects; since, as we have said before, the system detects the movement if the object is placed in the detection zone correctly, which was not always the case with these people. That is why we excluded people on level 4 of the study, and the results are shown in Table 5. The number of detection errors was drastically reduced for the group consisting of people on levels 1, 2, and 3. The total percentage of errors was 3.24, which is quite acceptable.

To complete this study, we also received the opinion of the trainers who helped us to carry out the experimentation of this research. Although we only had the information provided by three trainers, his opinion has been very valuable. In their opinion, the greatest contribution of this system is the possibility of tracing object movements on the worktable and the ability to record it for future analyze; or even that information can be served over the Internet in real time. Therefore, and in the opinion of trainers, LAoCA enables control and tracking capabilities; but as demonstrated by experimentation, it must be improved. These future improvements, further LAoCA versions, and experiments will be discussed in next section.

5.3. Is LAoCA Accepted by the Final Users? In this section, we present a study of the satisfaction perceived by the final users. In this research, there are two types of end users: participants and trainers. From participants, we wanted to find out two issues mainly: satisfaction and stress using LAoCA. The satisfaction was asked directly by the trainers to each of the participants after they finished the tasks, and the

stress level was analyzed by the trainers for each of the participants and agreed among them.

As Figure 14 shows the level of satisfaction is quite high, with an average of 4.125. This is because the participants took the experimentation as a game, since the worktable is quite attractive for the participants with four areas of detection of very striking colors, as well as the objects of detection that had different shapes and sizes. The participants left very enthusiastic and satisfied of the experimentation. However, during the experimentation, the participants suffered high degrees of stress as we can see in Figure 14; the mean of the stress level was 3.125. The participants did the experimentation one by one, in a room with three trainers observing the actions performed with the worktable and with several computers; as will be seen in Conclusions, it is recommended to perform this type of experimentation in a more environment-friendly manner for the participants. In addition, most of the participants presented their displeasure with the cyber-glove, which increased their level of stress and made them lose the attention of the tasks to be performed.

Also in this section, we present a study of the subjective perception of our approach from the point of view of trainers who participated in the experimentation. Because there were only three trainers who participated in the experimentation, we have sought the collaboration of seven other people to conduct a relevant survey about their opinion of our approach. These seven new people surveyed are specialists in education issues and described the qualities of LAoCA as they did a small test. Figure 15 shows a brief summary of the subjective evaluation of the ten specialists on three specific issues: utility, relevance, and satisfaction of our approach.

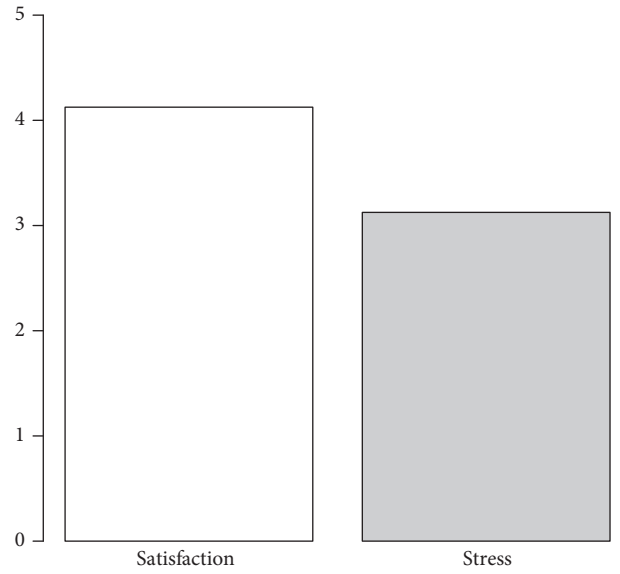


Figure 14: Evaluation of satisfaction and stress of the participants.

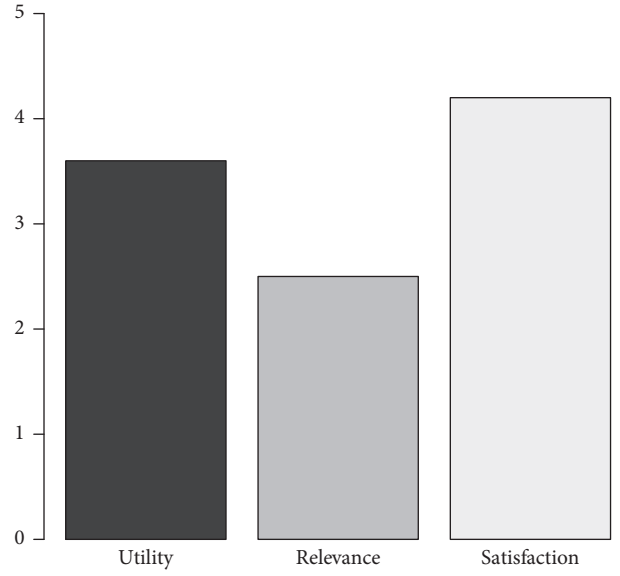


Figure 15: Evaluation of utility, relevance, and satisfaction of the trainers.

6. Conclusions and Future Work

In this research work, we present the second version of an approach for improving learning tasks for mentally handicapped people based on cyber-physical systems. The experimentation was carried out with the help and collaboration of the PRODIS Foundation, one of the most important foundations for mentally handicapped people in Spain. This research paper is an extended version of our previous work [6] presented in 10th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2016). Our approach is called LAoCA, and it is composed mainly of three elements: a worktable, a cyber-globe, and a software for modeling and executing workflows. LAoCA was built following main techniques on cyber-physical systems, ambient intelligence, and Internet of things. In order to validate our proposal, we designed an experimentation involving sixteen people with mental disabilities and their trainers. This experimentation tries to answer three research questions presented in Introduction. The aim of the experiment, proposed in this research work, was to test whether the experiment participants improved their learning of tasks using our proposal. For this, they had to repeat a sequence of movements with the tagged objects on the worktable. We will use research questions in order to guide this section, and at the end, we will present a set of lessons learned after experimentation.

6.1. Can LAoCA Facilitate Learning for Mentally Handicapped People? Studying all data from the experimentation indicates that there is no significant improvement in participants learning as all levels of disability are studied in the same data set. Analyzing the data in more depth, it can be observed that not all the people are able to improve their learning as it depends on the level of disability. Excluding people on levels 3 and 4 of the experiment, we can see that there is an improvement in learning for people on levels 1 and 2 (i.e., people with less disability). For people on levels 1 and 2, using a classical training, participants had trouble completing the tasks; however, using LAoCA, they performed less errors. For this reason, we can assume that an AmI system based on CPS techniques can be useful for

learning tasks for mentally handicapped people, but always bearing in mind that it is not easily applicable for people with a high level of disability. We also found a significant improvement for people on level 2 with easy task, but the statistical test had very little power because the sample is very small. This could be a point for further investigation. Therefore, the first conclusion we can draw is that not all people with mental disabilities are apt to use such devices; it is very important to know how to choose the set of people to work with, analyzing their level of disability and their ability to adapt to these devices. A possible future work would be to analyze how learning improves with people without mental disabilities, children, the elderly, etc.

6.2. Does It Improve Control over the Performed Tasks? To answer this research question, two sources of information have been studied: the first was a study about the detection failure from the sensors of LAoCA and the second was a survey of trainers about the benefits of it. Our approach sometimes does not detect the movements of tagged objects because the object is not placed in the center of the detection zone. We measured the number of detection failures and obtained that 8.33% of the movements are not detected correctly by LAoCA, which is not an acceptable value. But we realized that people on level 4 produced many detection errors, so we removed them from the study. In this way, the percentage of detection errors was reduced to 3.24. The conclusion that can be drawn reinforces that of the previous research question: not all people with disabilities can work with this type of device. The conclusion of the trainers is that the worktable is a good solution for control tasks and traceability and allows to control ambient assisted living tasks remotely, analyze the actions carried out a posteriori, and could even be developed an intelligent system that recognizes patterns in order to anticipate to critical situations.

6.3. Is LAoCA Accepted by the Final Users? In this section, we studied the level of satisfaction of the two groups of people that we consider the end users: the participants and the trainers. From the point of view of the participants, the level of satisfaction has been very high since the work done at the table was interpreted as a game. However, the participants according to their coaches suffered a lot of stress coming from a lack of familiarity with the participants. For future experiments, we propose two solutions: the first one is to create a familiar atmosphere with participants and the second consists to introduce new sensors to measure stress, blood pressure, sweat, etc. With these new sensors, we could know if the participant is suffering stress and we could stop the experimentation. From the point of view of the trainers, the utility, relevance, and satisfaction were studied, resulting in very high values for all parameters.

6.4. Lessons Learned after the Experiment Execution. After the execution of the experimentation with mentally handicapped people, we have drawn several conclusions that we would like to emphasize in this section. Regarding the cyberglove, it was an element of stress and loss of attention in most participants. We realized that something that is so attached to the body is not accepted by participants, generating doubts and rejection. That is why we recommend the use of elements that are as seamless as possible. As future work, we propose the development of a bracelet placed on the wrist able to detect objects caught with that hand. With respect to the worktable, the participants were distracted by the way it was built. For example, we used rivets to join the different pieces that make up the box. Those rivets were a source of distraction for the participants; most of them touched the rivets one by one until they discovered that they were not buttons.

In short and as a general conclusion, the system may be useful for people with a medium or low disability and for people who are not affected by other people or the environment. If the disability is more acute, it would require the presence of one or more evaluators. However, for people with a high disability or high concentration problems, this type of systems will not be useful since the system itself would be a source of distraction.

It is proposed as possible future work for the validation of each of the characteristics of the system proposed in this research work.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research has received funding from the Ministry of Economy and Competitiveness through the INPAINK (RTC-2016-4881-7) and SEMOLA (TEC2015-68284-R) projects, this work was also partially supported by the Autonomous Region of Madrid through program MOSI-AGIL-CM (grant P2013/ICE-3019, co-funded by EU Structural Funds FSE and FEDER).

References

[1] E. A. Lee, "Cyber-physical systems-are computing foundations adequate?," in *Proceedings of the Position Paper for NSF Workshop on Cyber-Physical Systems: Research Motivation, Techniques and Roadmap*, vol. 2, Austin, TX, USA, 2006.

[2] B. Bordel, R. Alcarria, T. Robles, and D. Martín, "Cyber–physical systems: extending pervasive sensing from control theory to the internet of things," *Pervasive and Mobile Computing*, vol. 40, pp. 156–184, 2017.

[3] B. Bordel Sánchez, R. Alcarria, D. Martín, and T. Robles, "TF4SM: a framework for developing traceability solutions in small manufacturing companies," *Sensors*, vol. 15, no. 11, pp. 29478–29510, 2015.

[4] G. Schirner, D. Erdogmus, K. Chowdhury, and T. Padir, "The future of human-in-the-loop cyber-physical systems," *Computer*, vol. 46, no. 1, pp. 36–45, 2013.

[5] D. Martín, R. Alcarria, Á. Sánchez-Picot, and T. Robles, "An ambient intelligence framework for end-user service

provisioning in a hospital pharmacy: a case study," *Journal of Medical Systems*, vol. 39, no. 10, p. 116, 2015.

[6] D. Martín, B. Bordel, R. Alcarria, Á. Sánchez-Picot, D. S. de Rivera, and T. Robles, "Improving learning tasks for mentally handicapped people using AmI environments based on cyber-physical systems," in *Proceedings of the 10th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2016)*, Part I 10, pp. 166–177, Springer, San Bartolomé de Tirajana, Gran Canaria, Spain, November 2016.

[7] D. J. Cook and S. K. Das, "How smart are our environments? An updated look at the state of the art," *Pervasive and Mobile Computing*, vol. 3, no. 2, pp. 53–73, 2007.

[8] National Institute of Standards and Technology, "CPS public working group presentation," 2015, http://www.nist.gov/el/upload/CPS-PWG-Kickoff-Webinar-Presentation-FINAL.PDF.

[9] D. Sanchez De Rivera, R. Alcarria, D. Martin de Andres, B. Bordel, and T. Robles, "An autonomous information device with e-paper display for personal environments," in *Proceedings of the 2016 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 139-140, IEEE, Las Vegas, NV, USA, 2016.

[10] D. Sánchez-de Rivera, D. Martín, R. Alcarria, B. Bordel, and T. Robles, "Towards a wireless and low-power infrastructure for representing information based on e-paper displays," *Sustainability*, vol. 9, no. 1, p. 76, 2017.

[11] International Electrotechnical Commission, *Enterprise-Control System Integration–Part 1: Models and Terminology*, ANSI/ISA-95.00.01-2010 (IEC 62264-1 Mod), American National Standards Institute and Instrumentation, Systems, and Automation Society, Research Triangle Park, NC, USA, 2010.

[12] ACR Advisory Group, "Automation and software expenditures for discrete industries," 2015, http://www.arcweb.com/market-studies/pages/automation-expenditures-for-discrete-industries.aspx.

[13] K. Thramboulidis and A. Zoupas, "Real-time java in control and automation: a model driven development approach," in *Proceedings of the 10th IEEE Conference on Emerging Technologies and Factory Automation (ETFA 2005)*, vol. 1, pp. 8–46, IEEE, Catania, Italy, 2005.

[14] M. Jimenez, F. Rosique, P. Sanchez, B. Alvarez, and A. Iborra, "Habitation: a domain-specific language for home automation," *IEEE Software*, vol. 26, no. 4, pp. 30–38, 2009.

[15] OPC Foundation, *OPC 2.1 Standard*, 1996, https://opcfoundation.org/.

[16] M. Philipose, K. P. Fishkin, M. Perkowitz et al., "Inferring activities from interactions with objects," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 50–57, 2004.

[17] C. Lee, M. Kim, J. Park, J. Oh, and K. Eom, "Design and implementation of the wireless RFID glove for life applications," *International Journal of Grid and Distributed Computing*, vol. 3, no. 3, pp. 41–52, 2010.

[18] K. P. Fishkin, M. Philipose, and A. Rea, "Hands-on RFID: wireless wearables for detecting use of objects," in *Proceedings of the Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, pp. 38–41, IEEE, Washington, DC, USA, 2005.

[19] J. Lee, B. Bagheri, and H. A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18–23, 2015.

[20] G. Schuh, T. Gartzen, T. Rodenhauser, and A. Marks, "Promoting work-based learning through INDUSTRY 4.0," *Procedia CIRP*, vol. 32, pp. 82–87, 2015.

[21] S. Katz, A. B. Ford, R. W. Moskowitz, B. A. Jackson, and M. W. Jaffe, "Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function," *JAMA*, vol. 185, no. 12, pp. 914–919, 1963.

[22] A. E. Majoros, B. C. Fredgren, P. R. Davies, and R. D. Kalinowski, "Data interface process with RFID data reader glove," US Patent 8482412, 2013.

[23] L. Atallah, B. Lo, R. King, and G. Z. Yang, "Sensor positioning for activity recognition using wearable accelerometers," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, no. 4, pp. 320–329, 2011.

[24] J. Hallam and V. Whiteley, "Interactive therapy gloves: reconnecting partners after a stroke," in *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA'11)*, pp. 989–994, ACM, Vancouver, BC, Canada, 2011.

[25] M. Konkel, V. Leung, B. Ullmer, and C. Hu, "Tagaboo: a collaborative children's game based upon wearable RFID technology," *Personal and Ubiquitous Computing*, vol. 8, no. 5, pp. 382–384, 2004.

[26] L. Escobedo, C. Ibarra, J. Hernandez, M. Alvelais, and M. Tentori, "Smart objects to support the discrimination training of children with autism," *Personal and Ubiquitous Computing*, vol. 18, no. 6, pp. 1485–1497, 2014.

[27] Lemelson MIT Program, "SignAloud: gloves that transliterate sign language into text and speech," 2016, https://www.youtube.com/watch?v=l01sdzJHCCM.

[28] P. S. López, J. Llinás, J. Amiel, C. Vidal, and J. Vilalta, "CAMDEX: a new psychogeriatric interview," *Actas Luso-Espanolas de Neurologia, Psiquiatria y Ciencias Afines*, vol. 18, no. 5, pp. 290–295, 1989.

Hindawi
Mobile Information Systems
Volume 2018, Article ID 4216172, 17 pages
https://doi.org/10.1155/2018/4216172

Research Article

A Fuzzy Logic-Based Personalized Method to Classify Perceived Exertion in Workplaces Using a Wearable Heart Rate Sensor

Pablo Pancardo, J. A. Hernández-Nolasco, and Francisco Acosta-Escalante

Informatics and Information System Academic Division, Juarez Autonomous University of Tabasco, Cunduacán, TAB, Mexico

Correspondence should be addressed to J. A. Hernández-Nolasco; adan.hernandez@ujat.mx

Received 28 July 2017; Accepted 12 November 2017; Published 22 January 2018

Academic Editor: Ljiljana Brankovic

Knowing the perceived exertion of workers during their physical activities facilitates the decision-making of supervisors regarding the worker allocation in the appropriate job, actions to prevent accidents, and reassignment of tasks, among others. However, although wearable heart rate sensors represent an effective way to capture perceived exertion, ergonomic methods are generic and they do not consider the diffuse nature of the ranges that classify the efforts. Personalized monitoring is needed to enable a real and efficient classification of perceived individual efforts. In this paper, we propose a heart rate-based personalized method to assess perceived exertion; our method uses fuzzy logic as an option to manage imprecision and uncertainty in involved variables. We applied some experiments to cleaning staff and obtained results that highlight the importance of a custom method to classify perceived exertion of people doing physical work.

1. Introduction

Advances in miniaturization, mobile communication, and sensor technologies make Mobile Health (mHealth) system development possible. mHealth is the intersection between Electronic Health (eHealth) and smartphone technologies. This means that the practice of eHealth is assisted by smartphones, which are used to capture, analyze, process, and transmit health-based information from sensors and other biomedical systems [1]. mHealth systems provide healthcare services with cost-effective, flexible, and efficient ways [2]. A mHealth system implemented on a mobile device enables a portable and nonobstructive solution; in addition, the computing and wireless capabilities allow real-time monitoring. This technology allows application deployment on mobile devices for continuous monitoring of people in order to determine, for example, the effort in their physical daily activities.

Humans possess a well-developed system for sensing the strain involved in physical effort. This is called perceived exertion (PE), which is the act of detecting and interpreting the sensations arising from the body during physical exertion [3]. Continuous measurement of physiological parameters in individuals while performing daily or labor activities allows health and well-being preservation or improvement.

Personal exertion estimation during labor activities has a particular interest, given that the effort to perform an activity is different for each person. Misallocation of an activity can affect a person's welfare and health. Workers may have risks associated with the disparity between high physical work demands and capacity/labor skills. These risks include musculoskeletal disorders, cardiovascular disease, prolonged absences, stress, burnout, and early retirements from the labor market [4]. Furthermore, physical strength assessment in ergonomics has additional benefits such as worker selection and placement and job design [5].

The estimation of workers' physical efforts in workplaces can be useful for allocation of employees in the appropriate position, adequacy of physical activities inherent to a job, prevention of accidents due to job demands, disease prevention related to physical demands, etc.

Generic methods known to estimate the physical effort do not take into account important physiological characteristics of individuals [6–8]. For many years, cardiac cost and

metabolic expenditure from physical labor are calculated using formulas and generic tables [6]. Physical exertion is then set, based on standards, such as the maximum heart rate (220-age). While in many cases this may be agile and convenient, it is not always true, as in the case of overweight or habituated people to perform an activity. It is necessary to develop methods that can provide higher accuracy for predicting energy consumption for a wide range of physical activities. This would allow a greater chance of being accurate on when to compare them to scientifically validated methods as doubly labeled water method [9].

Most available solutions for health monitoring offer a generalized physiological measurement, that is, by reference to generic formulas or tables that are not customized to individuals [7, 10]. Many other solutions are focused on predefined activities such as walking and running without considering physiological parameters of each person, giving results that are not clearly differentiated [9, 11].

In [12], a method based on personalized maximum heart rate was proposed as an extension of Chamoux method. This method allows continuous monitoring effort, taking into account the particular physical condition of each person by measuring the heart rate. The goodness of this proposal was evaluated in [13] through a comparative study with the other two methods (original Chamoux and Borg) [10, 14]. It can be stated from these results that the heart rate reflects health conditions (sick, tired, and acclimatized), but to our knowledge, this has not been proven objectively and formally, it can also be stated that the personalized maximum heart rate method allows a better result distribution than that obtained in previous works. However, these results do not consider the habituation to specific work that a person may have, nor the perception of experts about the nature of effort levels.

In this work, we propose a method considering both of these important factors in personalized effort evaluation: the habituation to perform a specific job and expert perception about the nature of effort levels assessment.

Expert knowledge refers to the estimates or judgments made by experts of the analysis and interrelations of problem's quantitative parameters. Usually, the expert knowledge must face situations of vagueness and imprecision. It is because it is complex or not possible to have a complete list of all variables involved in the problem domain. That is, there is no exhaustive list of all factors to take into account for the problem domain. Even knowing all the variables, it can be difficult to obtain concrete data. In addition, this information may be incomplete or even erroneous [15].

Expert perception of effort level is needed because of the nature of effort values reported in previous works based on relative cardiac cost (RCC), which is defined into a rank of values, RCC = [0, 69], and grouped in sets of 10 values—these sets are named as intense = [60, 69], heavy = [50, 59], slightly heavy = [40, 49], and so on [14]. The problem is when we have a value, let us say, RCC = 49, which is considered as slightly heavy, but which could be considered as heavy instead. In order to better define the effort magnitude, we propose to consider RCC sets as diffuse ones, given that there is a gradual progression of values from one set to the next, which allows us to define the membership degree of values to each set.

The habituation of a person to perform a specific activity is needed, because we must consider if this person has the skills needed to execute the activity with a good performance, that is, good performance in the execution of an activity depends more on habituation than on other factors as good physical condition. The habituation of a person to the execution of job activities affects the job after assignment in an important way.

2. State of the Art

For many years, cardiac cost and metabolic expenditure from physical labor are calculated using formulas and generic tables [6]. The use of a custom method becomes more important when monitoring physical activities that require a lot of effort, such as heavy lifting, since such activities are those that can compromise the welfare and health of workers [16].

As established in the ISO 8996 standard [17] for estimating metabolic cost, the use of the heart rate is an option that provides an estimation of effort with a margin of error as plus or minus 10 percent. This method of analysis is surpassed only by custom measurements that require the use of specialized equipment commonly available in laboratories. The latter very precise methods are equipment of indirect calorimetry (oxygen consumption test using a mask) and doubly labeled water (water consumption and urine analysis).

Measuring the heart rate is a valid option to estimate the effort which represents a work activity for an individual, although some limitations must be considered [18]. It is also important to consider that there are other factors influencing significantly, such as environmental conditions (temperature and humidity), weight, age, acclimation, mental stress, and personality [19].

2.1. Fuzzy Logic. A fuzzy logic provides an inference mechanism that allows us to simulate human reasoning into knowledge-based systems. The theory of fuzzy logic provides a mathematical framework that allows modeling the uncertainty of human cognitive processes in a way that can be treatable by a computer [15].

In accordance with [20], two important reasons to employ fuzzy logic are (1) data obtained from sensors measurements could be imprecise and imperfect and (2) fuzzy logic can deal with imprecision and uncertainty due to its properties of performance and intelligibility necessary for the classification process.

2.1.1. Fuzzy Logic Steps. Many solutions of real-world problems require dealing with inaccurate and imprecise data. Humans are able to solve these problems because they make use of cognition but also make use of fuzzy judgments and reasoning. Diffuse classification techniques have the advantage that require a soft decision, that is, a value that describes the degree to which an element belongs to a class. Instead of a hard decision, where one must say precisely whether an element belongs to a class or not, fuzzy logic is a very attractive field within artificial intelligence because it

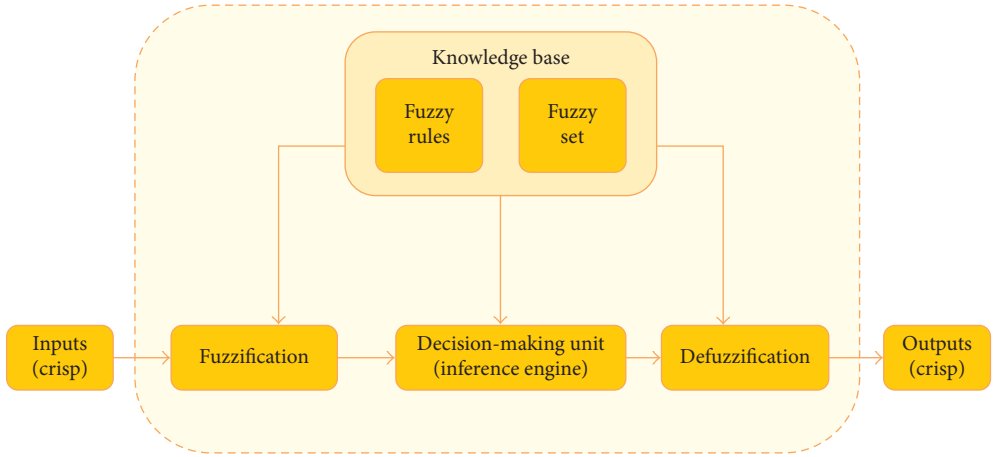


Figure 1: Fuzzy inference system steps.

is based on natural language. That is, it allows us to use linguistic terms to describe problems in a natural way. It does not use terms of relations between precise numerical values.

A fuzzy set can be defined as a set without clear and defined boundaries, in which elements that it contains can have a certain degree of membership ranging from total membership (value 1) to nonmembership (value 0). From this perspective, conventional sets (or crisp sets) can be seen as a particular case of fuzzy sets, a diffuse set that only admits two degrees of membership (one and zero).

Therefore, a diffuse set extends a standard set allowing degrees of membership of an element to the set, measured by the real numbers in the interval [0; 1]. If X is the universe of discourse and its elements are denoted by x, then a fuzzy set A on X is defined as a set of ordered pairs $(x, \mu_A(x))$ such that

$$A = \left\{ x, \frac{\mu_A(x)}{x}, \quad 0 \le \mu_A(x) \le 1 \right\}, \tag{1}$$

where $\mu_A(x)$ in (1) is the membership function of each x in A. In contrast to classical logic where the membership function $\mu_A(x)$ of an element x belonging to a set A could take only two values: $\mu_A(x) = 1$ if $x \in A$ or $\mu_A(x) = 0$ if $x \notin A$, fuzzy logic introduces the concept of membership degree of an element x to a set A and $\mu_A(x) \in [0; 1\ ]$; here we speak about the truth value.

A typical fuzzy logic inference system has four components: the fuzzification, the knowledge base (rules and fuzzy sets), the inference engine, and the defuzzification [15]. Figure 1 shows those main fuzzy inference system steps.

2.1.2. Fuzzification. The first step in fuzzy logic is to take the measured data (crisp data) and determine membership degree of these inputs to associated fuzzy sets. It is done by giving a value to each variable to a membership function set. Naturally, crisp value will be limited to the universe of discourse. Membership functions take different shapes. The two most common functions are triangular and trapezoidal. A triangular membership function with straight lines can formally be defined as follows:

$$\Lambda(x, a, b, c) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ \dfrac{c-x}{c-b}, & b \le x \le c \\ 0, & x \ge c. \end{cases} \tag{2}$$

Trapezoidal function is shown in the following equation:

$$f(x, a, b, c, d) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \\ \dfrac{d-x}{d-c}, & c \le x \le d \\ 0, & x \ge d. \end{cases} \tag{3}$$

A Gaussian membership function with the parameters m and σ to control the center and width of the function is defined by

$$G(x, m, \sigma) = e^{-(x-m)^2/2\sigma^2}. \tag{4}$$

The generalized Bell function that depends on three parameters a, b, and c is given by

$$f(x, a, b, c) = \frac{1}{1 + ||(x-c)/a||^{2b}}. \tag{5}$$

Other membership functions are sigmoid-shaped functions and delta functions (single functions). Selecting the membership function will depend on the nature of the problem, the type of data, and the experimental results. A knowledge expert is important to decide which shape will be used.

2.1.3. Knowledge Base (Rules and Fuzzy Sets). Rules are constructed from linguistic variables. Rules are structured in a IF-THEN format. The IF part of the rule is the antecedent and the THEN part of the rule is the consequent. These variables take on the fuzzy values that are represented as words and modeled as fuzzy subsets of an appropriate domain. There are several types of fuzzy rules, we mention only the two main rules:

(i) *Mamdani rules* [21]: These rules are of the following form: if x_1 is A_1, x_2 is A_2, . . ., x_p is A_p, then y_1 is C_1, y_2 is C_2, . . ., y_p is C_p, where A_i and C_i are fuzzy sets that define the partition space. The conclusion of a Mamdani rule is a fuzzy set. It uses the algebraic product and the maximum as T-norm and S-norm, respectively, but there are many variations by using other operators.

(ii) *Takagi/Sugeno rules* [21]: These rules are of the following form: if x_1 is A_1, x_2 is A_2, . . ., x_p is A_p, then $y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_px_p$. In the Sugeno model, the conclusion is numerical. The rules' aggregation is in fact the weighted sum of rules' outputs.

2.1.4. Inference Engine. The fuzzy inference system uses fuzzy equivalents of logical AND, OR, and NOT operations to build up fuzzy logic rules. An inference engine operates on rules to evaluate them. Inference engine takes inputs and applies them to the antecedent part of the rule. If a rule has multiple antecedents, then logical AND, OR, and NOT operations are used to obtain a unique value representing evaluation result. This result (truth value) is applied to the consequent part. The outputs are then added. It is the process of unification of the outputs of all rules, that is, the membership functions of all consequent previously trimmed or scaled outputs are combined, to obtain a single fuzzy set for each output variable.

2.1.5. Defuzzification. The final step of a fuzzy logic system consists of transforming the fuzzy variables obtained by the fuzzy logic rules into crisp values again that can then be used to take a decision or perform some action. There exists different defuzzification methods: centroid of area (COA), bisector of area (BOA), mean of maximum (MOM), smallest of maximum (SOM), and largest of maximum (LOM). In our system, we used COA, and the following equation illustrates it:

$$Z_{\text{COA}} = \frac{\sum_{i=1}^{n} \mu_A(x_i) x_i}{\sum_{i=1}^{n} \mu_A(x_i)}. \quad (6)$$

2.2. Habituation. Habituation is a form of learning in which an organism decreases or ceases its responses to a stimulus after repeated presentations [22]. In perceived exertion, context is about how much a person has repeated a physical activity. Habituation as a state of training affects the heart rate [23]. In labor context, it refers to how frequently workers perform a specific physical activity related to their job.

Habituation to the performance of physical work activities is important because a person not being accustomed to perform a specific physical activity has a perceived effort of about twenty percent higher than a person accustomed to performing such an activity [24, 25].

Several methods have been used to quantify workload, including questionnaires, diaries, physiological monitoring, and direct observation [23]; in this sense, direct observation method can be considered to determine habituation, based on intensity and frequency (workload) of individual daily activities' performance.

3. Related Work

There are several works related to the proposal that we present, but none with the approach (personal perceived exertion), combination of factors (habituation, relative cardiac cost, and degree of membership to a fuzzy group), and application domain (prevention of labor accident risks due to workload fatigue) that is proposed.

The first group contains studies in workplaces oriented to estimate energy expenditure or activity recognition using technological devices. For example, Hwang et al. [26] proposed a measurement approach in energy estimation field. It is expected to provide in-depth understanding and continuous monitoring of worker's physical demands from construction tasks. Their solution was to use the heart rate (HR) to estimate EE according to a linear relationship between HR and EE. Their proposal was to achieve reliable field EE measurement through automatic action recognition using an embedded accelerometer and applying HR-EE relationships for corresponding actions with acceptable HR monitoring accuracy.

Hwang's proposal is based on identifying physical activities, which to date is limited to certain activities such as walking, running, and climbing stairs. That is, we could not identify any physical activity derived from a job; this makes Hwang's proposal not suitable for any type of work where physical activities are performed. On the contrary, our proposal focuses on identifying the personal physical effort involved in the work activity, without needing to identify which is the activity that the worker performs. Another example is shown in [27]; in this case, authors estimate and try to predict energy expenditure predictions based on the heart rate. On the contrary, we are compelled to estimate perceived exertion.

The second group contains works aimed at preserving health at work. For example, Migliaccio et al. [28] used sensors to monitor physical bends performed by construction workers, so it is identified that those physical activities can be risky to health. In this experiment, a heart monitor was used to detect high heart rates which were directly associated with a subject carrying a load. Fusing heart rate data and posture data provided the capability of differentiating safe from unsafe material-handling activities. The main objective of this research was to assist future decision makers in designing ergonomically safe and healthy

work environments. Migliaccio's work focuses on detecting high levels of heart rate and unsafe postures, but the proposal is not personalized.

Aryal et al. [29] present a method for real-time monitoring of physical fatigue in construction workers using heart rate monitoring and infrared temperature sensors. Boosted tree classifiers were trained using the features extracted from the heart rate and temperature sensor signals and used to predict the level of physical fatigue from 12 participants. The study lacks a personalized classification of effort since it uses the Borg scale, which is extremely generic and does not contribute to the personalized detection of the efforts. There is a non-personalized classification because during physical activity, relative effort regarding resting heart rate and personalized maximum heart rate is not considered.

The third group contains those researches based on the fuzzy logic. The fuzzy logic-based tool for modeling human sensitivity to thermal sensation developed by Shimizu and Jindo [30] demonstrated that membership functions capture the ambiguity of classes to categorize thermal sensations. In the same sense, in [31] the theory of fuzzy sets and systems was applied to assess perceived workload involved in manual lifting tasks. In [32] a fuzzy logic-based risk assessment framework to evaluate physiological parameters is proposed; this model is used to avoid emergency situations during sport activity; however, personalized heart rate thresholds used in this proposal are based on generic values [10] from runners and triathletes.

These results support that our hypothesis about the fuzzy logic is convenient for classifying humans' effort perceptions. Additionally, no proposal considers the impact of habituation to work on physical effort, nor the degree of membership that has a cardiac cost value to a defined effort class.

3.1. Heart Rate-Based Methods to Estimate Physical Effort. In this paper, we use methods based on the heart rate because this type of parameter has a 90% accuracy in estimating physical efforts, as it is stated in safety and health standards at work [17].

There are several methods that rely on measuring the heart rate to establish which is the physical effort that a work activity can represent for people [33]. We selected two of them: the Borg rating scale of exertion [10] and the Chamoux method [14].

The Borg scale is widely known and applied in sport and medical domains; it is generic and based on a table where, if a person has a certain value of heart rate, then it has a certain level of effort, and it is called rating of perceived exertion. In Table 1, the Borg scale shows 14 (6 to 20) values grouped in six categories.

In order to interpret the Borg scale, the numbers in the left column correspond to the number of beats of one person during physical activity divided by 10, and the corresponding value in the right column is the perceived exertion (level of effort); for example, if a worker has 110 beats per minute, the level on the scale is 11 and it belongs to slight effort. In this method, it is assumed that the maximum heart rate of a person is 220 minus his/her age in years. A real effort test is not required; therefore, it is a generic value.

Table 1: Borg's scale.

Scale	Description
6	No exertion
7	—
8	—
9	—
10	—
11	Light
12	—
13	Somewhat hard
14	—
15	Hard (heavy)
16	—
17	Very hard
18	—
19	—
20	Maximal exertion

Table 2: Different levels of effort for RCC under Chamoux.

RCC	RCC level	Effort
0–9	RCC1	Very light
10–19	RCC2	Light
20–29	RCC3	Slightly moderate
30–39	RCC4	Moderate
40–49	RCC5	Slightly heavy
50–59	RCC6	Heavy
60–69	RCC7	Intense

Otherwise, Chamoux [14] proposes a lesser-known method, and as far as we know, it is not frequently used. This method requires to measure resting and the maximum heart rate for each person, taking into account several physiological parameters.

The method consists of two steps to estimate the physical effort. This first step is to obtain labor activity's absolute cardiac cost (ACC), which is obtained using the average cardiac frequency (ACF) and the resting cardiac frequency (RCF) for a person at every moment. ACF is obtained from the average value of the frequency of the worker during a day of conventional job. RCF is obtained after a person has slept (8 hours) and is resting.

ACC is obtained by subtracting the resting cardiac frequency (RCF) from the average cardiac frequency (ACF), as shown in the following formula [14]:

$$\mathrm{ACC} = \mathrm{ACF} - \mathrm{RCF}. \tag{7}$$

The second step is to compute the relative cardiac cost (RCC). Therefore, we should get theoretical maximum cardiac frequency (TMCF). Conventionally, the TMCF value is obtained by subtracting the person's age in years from 220. The formula for RCC is as follows [14]:

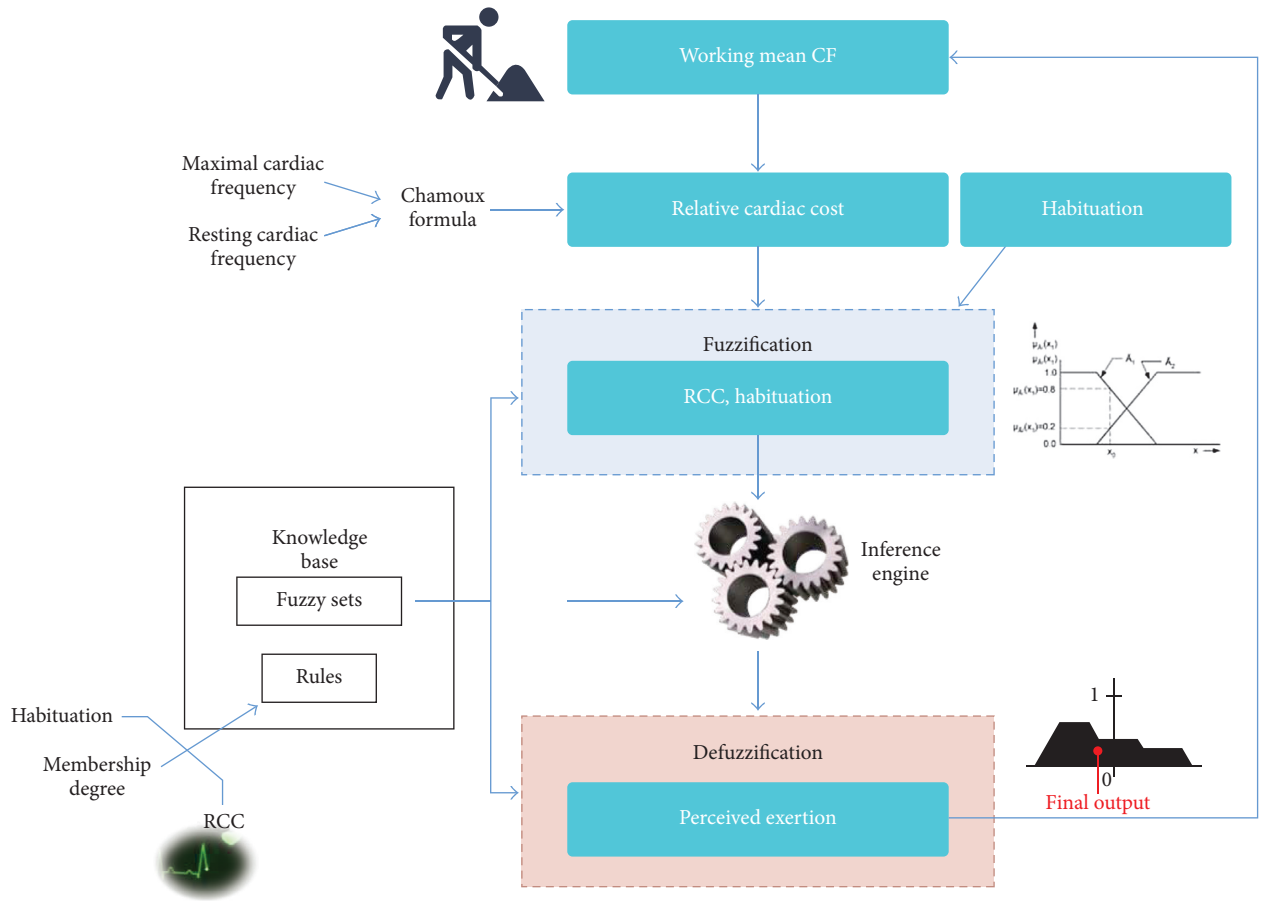


Figure 2: Proposed method (FPC).

$$\text{RCC} = \left(\frac{[\text{ACC} * 100]}{[\text{TMCF} - \text{RCF}]} \right). \quad (8)$$

Effort levels for a worker according to the method of Chamoux are shown in Table 2.

4. Materials and Method

We used a Basis B1 Fitness Wristband as a heart rate monitor, an Omron Sphygmomanometer Model HEM-742INT, and a common stethoscope. We used a treadmill mark BH Fitness Model Prisma M60 to measure personalized maximal cardiac frequency. It was operated without inclination. The prototype to estimate perceived exertion was developed with the Java 6.0 language using the ADT tool v22.3.0-887826. The prototype was implemented over a Samsung Galaxy S4, an Android 4.2.2 (Jelly Bean) Operation System, an octa-core chipset, and a 1.6 GHz Quad + 1.2 GHz Quad CPU.

We propose a method based on the method of Chamoux as it was explicitly created for the work environment, while the Borg method is used in sports. The fuzzy personalized Chamoux-based method (FPC) we propose is illustrated in Figure 2. Our fuzzy inference method is Mamdani type.

The first step of the proposed FPC method is taking cardiac frequency at rest, personalized maximal cardiac frequency, and habituation level. As we decided to customize the Chamoux method, that is, obtaining the value of TMCF parameter for each person, we required each user to perform a maximal exercise stress test using an electric treadmill and we took the value of the heart rate as their TMCF. We refer to this as a personalized Chamoux method [12]. Habituation value was assigned considering how frequent and experienced the user is about a specific labor physical activity.

Users will carry a wearable heart rate monitor to have a continuous monitoring of cardiac frequency during labor activity. From this monitoring, we obtain media cardiac frequency per minute. Having these data, we apply the Chamoux formula, using as maximum cardiac frequency, the personalized value that was obtained during the test with the treadmill. Chamoux formula gives us relative cardiac costs (RCCs) for each worker.

Later, RCC variables were assigned to fuzzy sets, as shown in Figure 3.

In accordance with knowledge obtained in [25], it was used as a condition that establishes that nonhabituated workers increase their perceived exertion by 20% for evaluated activity (sweeping, cleaning windows, and stacking chairs). Another condition was implemented for moderately habituated workers; in this case, their PE was increased 10%. For habituated workers, there is no need to increase

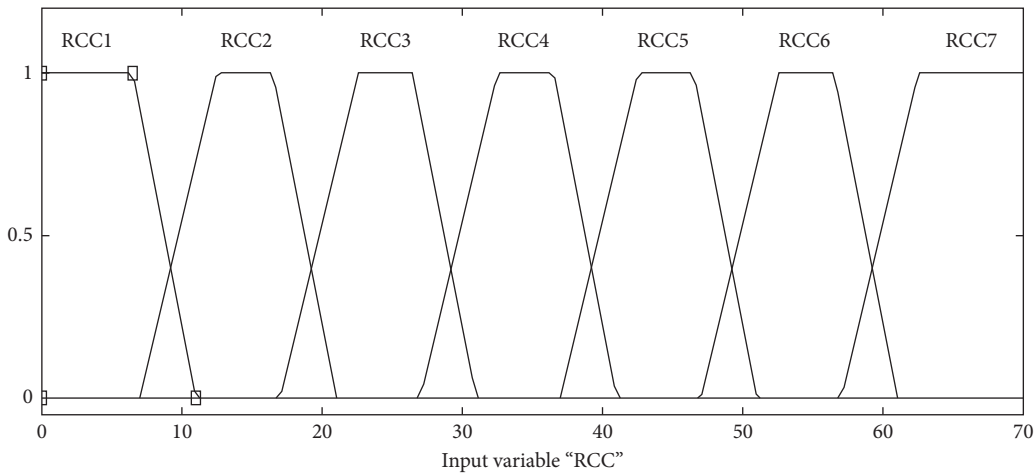


Figure 3: RCC membership function.

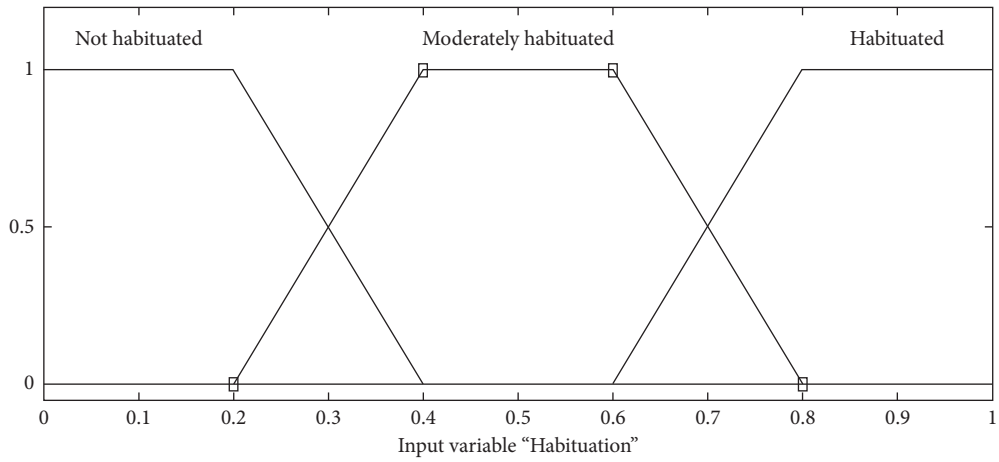


Figure 4: Habituation function.

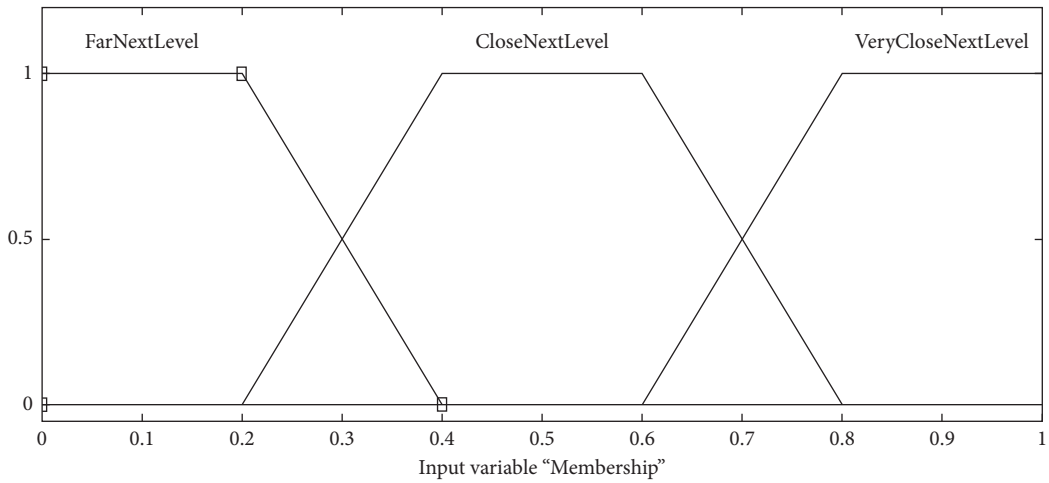


Figure 5: Level of membership function.

If (RCC is RCC1) and (Habituation is habituated) then (perceived Exertion is PE1)

If (RCC is RCC1) and (Habituation is ModeratelyHabituated) and (Membership is FarNextLevel) then (Perceived Exertion is PE1)

If (RCC is RCC1) and (Habituation is ModeratelyHabituated) and (Membership is CloseNextLevel) then (Perceived Exertion is PE1)

If (RCC is RCC1) and (Habituation is ModeratelyHabituated) and (Membership is VeryCloseNextLevel) then (Perceived Exertion is PE2)

If (RCC is RCC1) and (Habituation is NotHabituated) and (Membership is FarNextLevel) then (Perceived Exertion is PE1)

If (RCC is RCC1) and (Habituation is NotHabituated) and (Membership is CloseNextLevel or Membership is VeryClose NextLevel) then (Perceived Exertion is PE2)

Figure 6: Rules for RCC1.

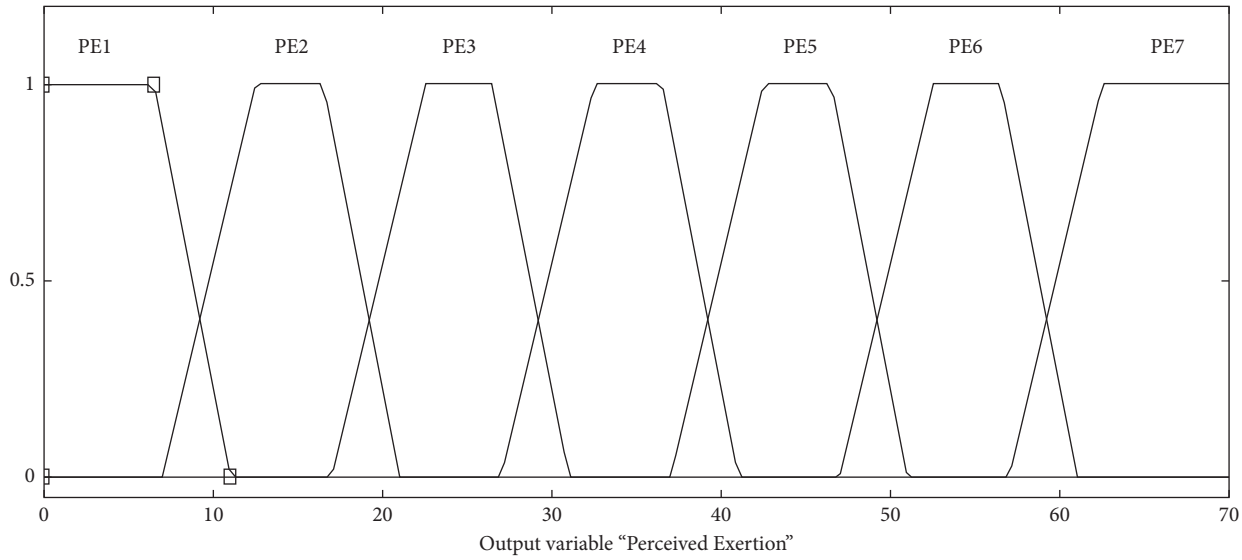


Figure 7: Output variable "perceived exertion."

(compensate) the perceived exertion. Membership function for habituation variable is shown in Figure 4.

Another criterion used to conform the rules is a variable called membership. This variable is used to define when a PE result must be located in the next level. That is, if a worker is moderately habituated and his membership variable is VeryCloseNextLevel, then his PE is upgraded to the next level of PE. If a worker is nonhabituated and his Membership variable equals to CloseNextLevel or VeryCloseNextLevel, then his PE result is upgraded to the next level of PE. Figure 5 illustrates the Membership function.

After that, rule base is constructed, rules are based on knowledge or experience in the domain and these are useful for the inference engine to carry out the process of defuzification. In our proposal, rules make use of the fuzzification of efforts scalar sets defined by Chamoux and habituation impact on the worker.

An extract of the rules used in the proposal is illustrated in Figure 6, specifically for the first level of relative cardiac costs (labeled as RCC1). For each one of the seven RCC levels [RCC1, RCC7] were built a group of rules similar to these.

Later, we use the inference engine fed by the defined rules, where the level of habituation and the degree of membership to relative cardiac costs are part of the rules. The last step is the defuzzification phase. The COA method was applied for defuzzification. The COA defuzzification method effectively calculates the best compromise between multiple output linguistic terms [15]. In this phase, we obtained PE, which is one of all possible outputs (linguistic variables), as shown in Figure 7.

5. Experiments

The tests were conducted on a university campus, and users were potential janitors and janitors. In the experiments, a group of 20 research participants conducted a series of work activities, and heart rate measurements were taken during those activities. These data sets were collected using a population of 20

Table 3: Participants' characteristics.

Subject	Genre	Age	BMI	Habituation level
Worker 1	Male	23	29.62	Habituated
Worker 2	Female	23	21.98	Habituated
Worker 3	Female	23	20.95	Moderately habituated
Worker 4	Male	22	24.72	Habituated
Worker 5	Male	23	28.27	Habituated
Worker 6	Male	23	20.03	Not habituated
Worker 7	Female	24	19.27	Not habituated
Worker 8	Female	33	31.24	Moderately habituated
Worker 9	Female	34	21.68	Habituated
Worker 10	Male	25	28.44	Habituated
Worker 11	Male	28	28.33	Habituated
Worker 12	Male	27	24.38	Habituated
Worker 13	Male	24	33.64	Not habituated
Worker 14	Female	28	29.36	Moderately habituated
Worker 15	Male	34	23.63	Not habituated
Worker 16	Female	36	24.35	Not habituated
Worker 17	Male	33	24.68	Habituated
Worker 18	Female	22	26.29	Not habituated
Worker 19	Female	36	30.44	Not habituated
Worker 20	Male	51	23.12	Habituated

Table 4: Activities in the experiments.

Activity	Description
Sweep the floor	One broom (1 kg) was used in this activity. A hallway (42 m long × 0.5 m width) was the area to sweep. The volunteers started in one corner of the hallway and swept in overlapping strokes in towards the end of the hallway.
Washing windows	The total dimensions of the window were 110 cm × 90 cm. The research participants started at the top and worked down the window. This activity was executed in indoor environments.
Stacking chairs	Placing the entire stack of chairs had a short walk away (3 m). This activity was done using iron chairs (7 kg). During these activities were created several stacks, each stack with 8 chairs; all experiments were done in an indoor hallway. Never were stacked more than ten chairs at a time.

participants; 11 male (28.4 ± 8.5 years, BMI 26.26 ± 3.77) and 9 female (28.7 ± 5.97 years, BMI 25.06 ± 4.45). Participants' characteristics are shown in Table 3.

Three physical activities were defined for every research participant. These activities are described in Table 4.

Personal characteristics and physical conditions (such as age, sex, acclimation, and physical condition) are the attributes that are indirectly reflected when we measure the maximal theoretical heart rate being their maximal personal effort for each user. Together with the heart rate at rest and individualized monitoring in real time during the execution of physical activities, they allow customized estimations. During analysis, these characteristics' results allow us to see that two people with similar characteristics do not necessarily perform the same effort to perform the same activity.

Heart rate was measured using an unobtrusive Basis B1 fitness tracker band. Basis' precision is enough to know how many beats per minute a user heart has. Basis B1 measures our blood pressure, steps, intensity and exertion of our workout, and sleep metrics. This device was placed on the wrist of each worker. The first activity was to sweep a floor using a broom, the second activity was to clean glass windows with a rag, and the last activity was stacking metal structure chairs.

Heart rate values used in all methods (Borg–Chamoux–personalized Chamoux–fuzzy personalized Chamoux) were the average heart rates during the activities.

Experiments related to the three labor activities are shown in Figure 8.

6. Results

In order to compare the resulting values of all methods tested, we made a mapping of the Borg's perceived exertion values (Table 1) with labels used in Chamoux method (Table 2). Scales 6-7 are no exertion (NE), 8-9 are very light (VL), 10-11 are light (L), 12 is slightly moderate (SM), 13 is moderate (M), 14 is slightly heavy (SH), 15-16 are heavy (H), and over 16 is intense (I).

A frequency analysis of results of perceived exertion of the participants obtained for each physical activity was included. In order to do this, we obtained some values describing the features of a collection of data from physical activities performed. For stacking chairs activity, Table 5 shows the number of users for each perceived exertion level grouped by the method.

In Table 5 we can see that the estimated perceived exertion of people using the Borg method is only two levels, the common Chamoux method classifies them into three levels, and the personalized Chamoux and our proposal (FPC) classify them into five levels. The Borg method classifies all people into very light (VL) and light (L); conventional Chamoux classifies 20% into SM, as it only takes into account the age of the people; personalized Chamoux distributes 60% of workers between SM and SH, this is because it takes into account personal maximum effort, in addition to the age of the individual; while the proposed method makes a small rearrangement of the number of people at every level, which results from applying the fuzzy logic for handling uncertainty membership groups and the effect of habituation variable. This indicates that our proposal has a better effort discrimination because of measuring their personal maximum effort, fuzzy sets without clear and defined boundaries, rules base, and habituation-level variable.

In Figure 9(a), the results of perceived exertion are scalar (diffuse for FPC), whereas in Figure 9(b) the results of perceived exertion are linguistic (crisp for FPC). The objective is that the decision-maker can appreciate not only the level of final perceived exertion obtained (after the whole process) but also the level of belonging to that level (scalar values). The same criterion applies for Figures 9(c) and 9(d).

(a) (b) (c)

Figure 8: Activities developed by participants: (a) sweeping floors, (b) cleaning windows, and (c) stacking chairs.

Table 5: Number of users for each effort level during stacking chair activity, grouped by method.

Perceived exertion	Borg	Chamoux	Personalized Chamoux	Fuzzy personalized Chamoux
VL	11	4	0	0
L	9	12	8	6
SM	0	4	6	6
M	0	0	4	5
SH	0	0	2	3

These results can be used in decision-making to preserve or improve the health and quality of life of the worker. This can be done by adjusting their work environment or by measuring physical performance based on their effort for a better allocation of their workload.

As we can see in Figures 9(a) and 9(b), a female worker who is moderately habituated to physical job maintained her perceived exertion level obtained using FPC with respect to the personalized Chamoux with sweeping and cleaning windows activities; however, while she was stacking chairs (which demands more physical effort), her perceived exertion level increased. With regard to not habituated male Worker 6 (Figures 9(c) and 9(d)), his perceived exertion level obtained using FPC increased with respect to the personalized Chamoux with all activities (sweeping, cleaning windows, and stacking chairs). All his FPC perceived exertion levels were higher than the personalized Chamoux perceived exertion levels. We attribute this behavior to the level of habituation. One objective of this proposal is to illustrate how the habituation factor impacts the perceived effort, and the proposal is not focused on an accurate calculation of physical effort or energy expenditure.

Figures 10 and 11 show personalized Chamoux and FPC methods to classify perceived exertions during sweeping.

As we can see, results clearly reflect different perceived exertion levels for individuals even though they perform the same activity. In Figure 10, the results of the FPC method are fuzzy, and in Figure 11, the results of the FPC method are defuzzified.

The statistical results allow us to know that when the activity is physically demanding (activity of stacking chairs), variance and standard deviation values are much higher. This shows that the many factors involved in the process of classifying perceived exertion are clearly reflected in the increase in cardiac frequency.

Figures 12 and 13 illustrate an activity that can be physically demanding if we are not habituated, and they show how the fuzzy personalized Chamoux method is more efficient for classifying individual perceived exertion, which is appreciated particularly for the activity of stacking chairs (Figures 12 and 13). Figure 12 can be very useful for a decision-maker to appreciate how a worker is being impacted for a specific physical activity.

All participants were directly observed during experiments to estimate their physical effort level. Additionally, they were asked about their perceived exertion just at the end of each activity. We obtained that perceived exertion classification using our proposed method is coincident (75% or higher) with respect to our direct observation and answers from participants.

Workers' personal perceived exertion (PPE) was a linguistic label (belonging to Table 2) that participants assigned

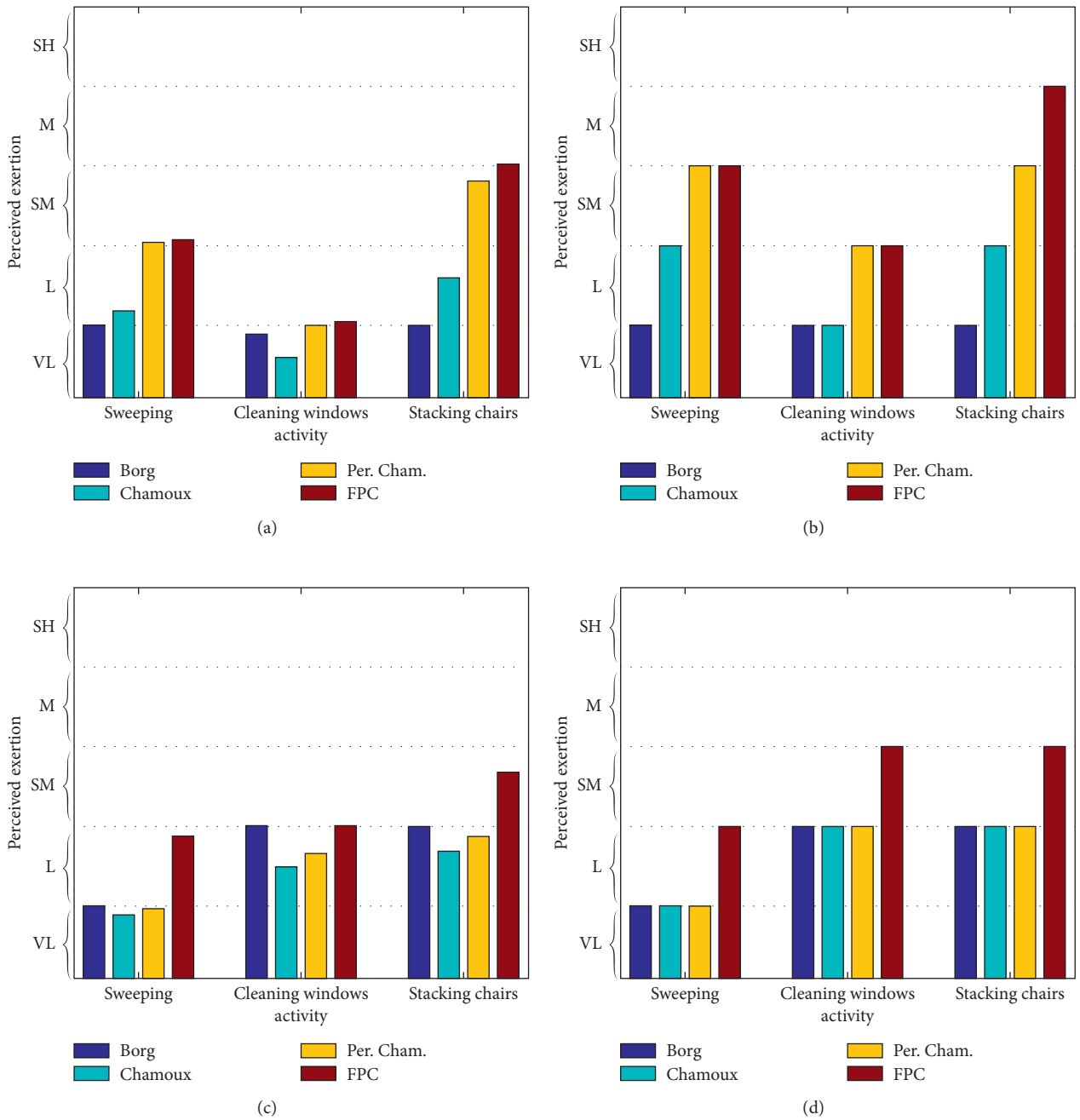


Figure 9: Perceived exertion: sweeping, cleaning windows, and stacking chairs. (a) Worker 3 (moderately habituated female—scalar values), (b) Worker 3 (moderately habituated female—linguistic values), (c) Worker 6 (not habituated male—scalar values), and (d) Worker 6 (not habituated male—linguistic values).

to each developed activity, representing their effort perception. We compare PPEs with FPC results to evaluate our proposal. Matching percentages were 75%, 75%, and 80% for sweeping, cleaning windows, and stacking chairs, respectively. As an example, in Figure 14, it is appreciated a comparison of workers' PPE versus linguistic outputs provided by our proposal for the sweeping activity.

We have designed a prototype for logging and informing users about their perceived exertion levels and historical records during activities. Figure 15 shows the prototype for Android devices.

The disk located at the bottom of the interface simulates a semaphore. The colors used represent the different levels of perceived exertion (from very light to intense). The green color represents the lowest levels of effort, and the red color represents the intense effort. The yellow color is used for moderate efforts. The purpose of the disk in the interface is that a given user can visualize in each moment and, in real time, the percentage of monitoring time that he/she has been in each level of perceived exertion, in accordance with our proposed method. Recording of perceived efforts is useful for rapid decision-making by the supervising manager, for

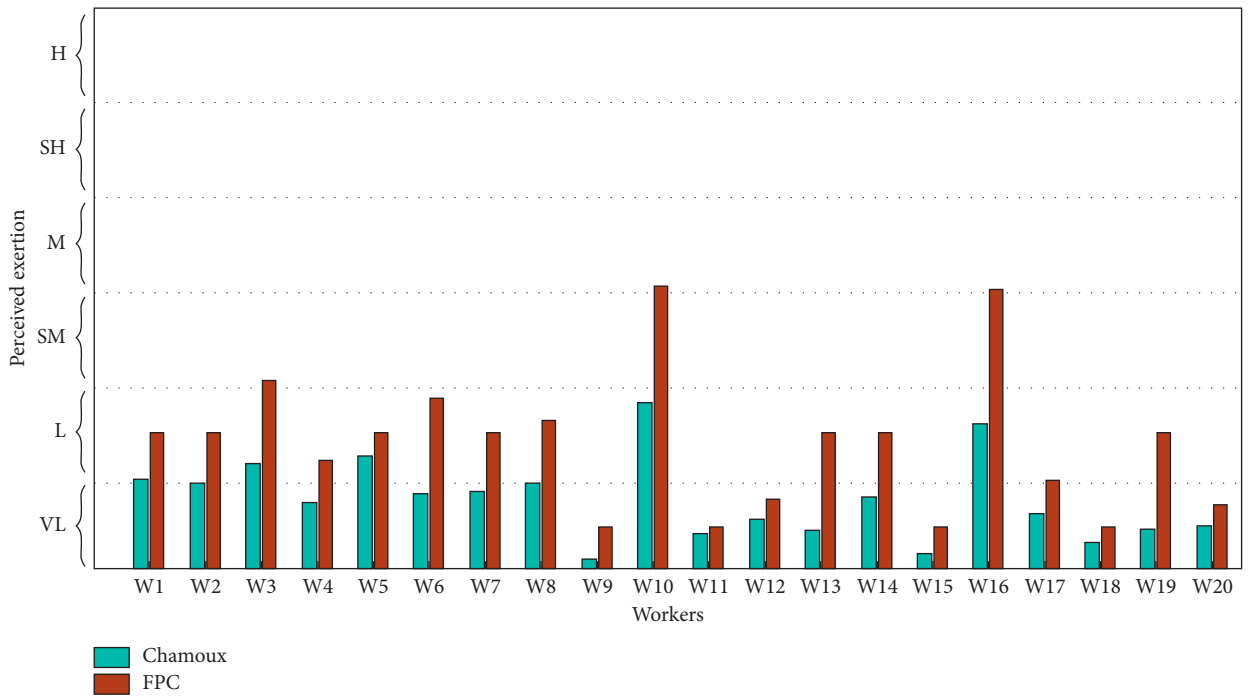


Figure 10: Sweeping activity (scalar results).

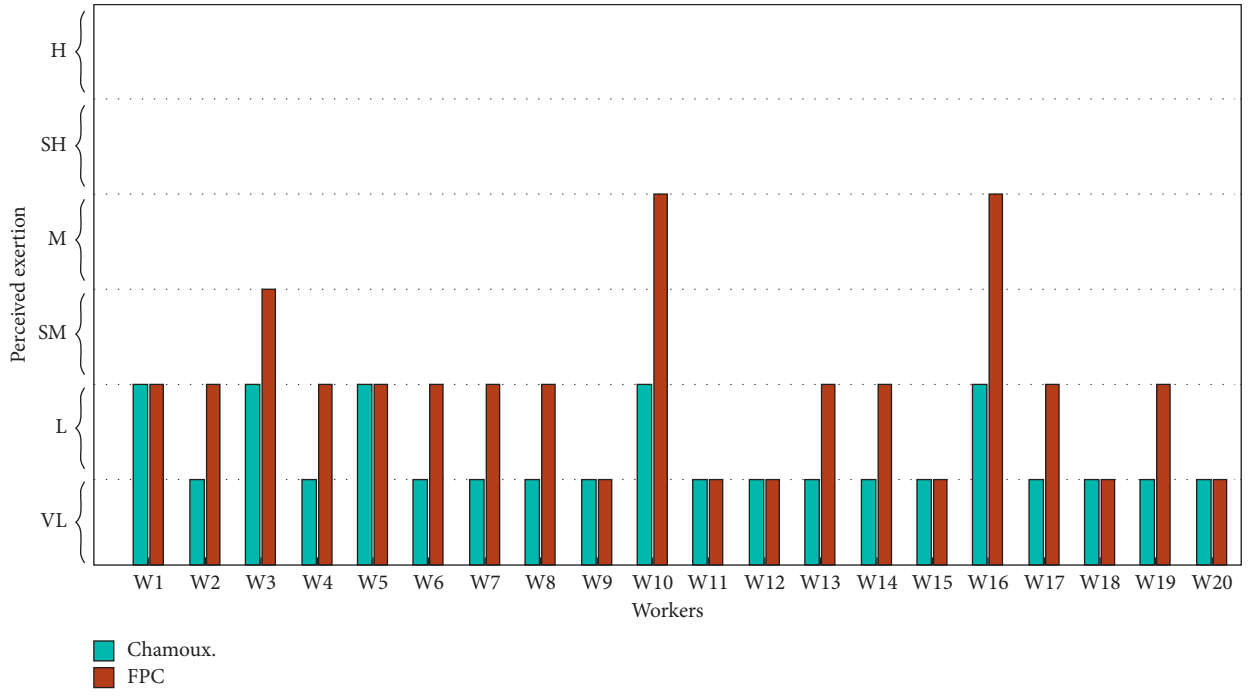


Figure 11: Sweeping activity (linguistic results).

example, for reassignment of tasks or scheduling of rest periods.

7. Discussion

The smartwatch presented several failures when data started to be captured, that is why a period of at least three minutes was monitored before the value of cardiac frequency was taken. However, as cardiac frequency values entered to formulas were averaged, some noise effects do not impact the results. We think that technological advances will let that future devices be more precise in sensing.

Considering habituation as a factor in the method of estimating perceived effort allows in turn improving

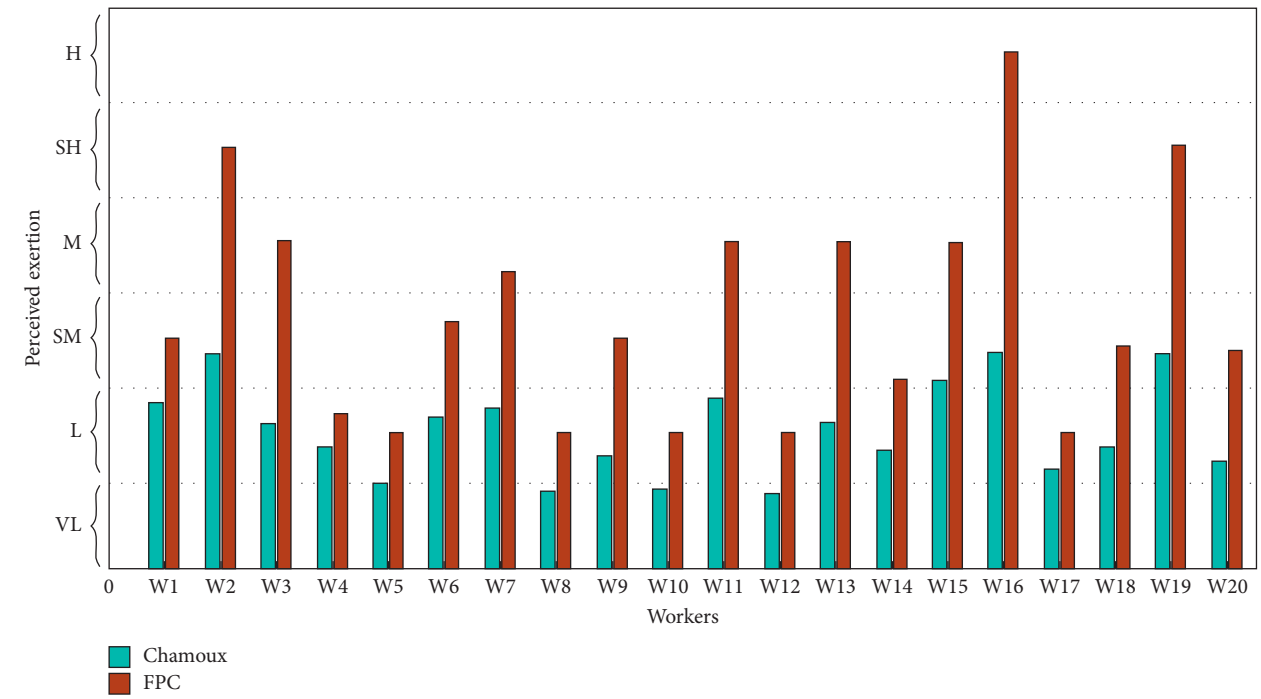


Figure 12: Stacking chair activity (scalar results).

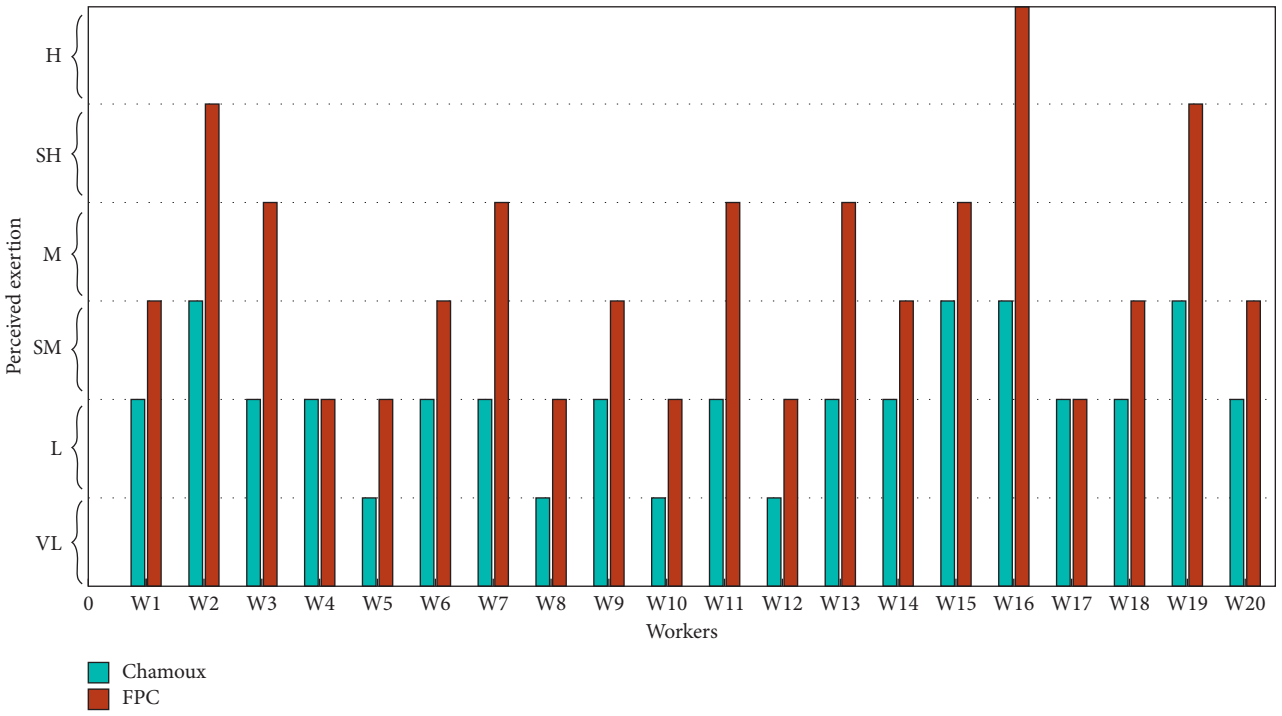


Figure 13: Stacking chair activity (linguistic results).

decision-making, for example, when a worker should be selected for a job with physical activities. Habituation is important because an unhabituated worker will have a greater perceived effort with respect to another worker who is accustomed, as studies reveal [24, 25].

Although recent proposals have determined rates in perceived exertion, these results do not consider individual factors related to worker experience or the physical activity performance. This is, for activities that cannot be controlled in terms of the intensity with which users perform them

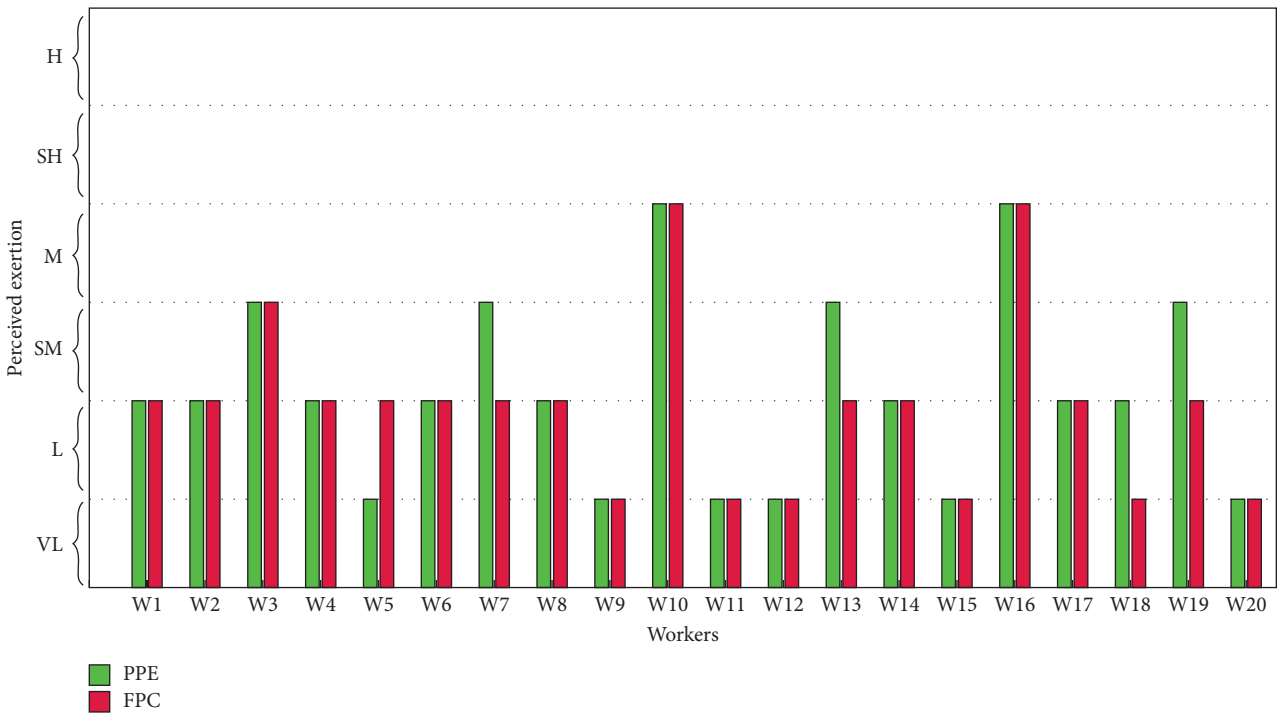


Figure 14: Sweeping activity (personal perceived exertion versus FPC method results).

(e.g., sweeping), each user executes them according to their personality, unlike activities performed on an electric treadmill where the speed at which they walk or run is controlled, so many studies only offer average energy expenditure values for daily activities. For this reason, in some cases for the same activity, a moderately habituated person may have a perceived exertion slightly greater than an unhabituated person.

Perceived exertion assessment in labor physical activities must consider factors such as physical condition, obesity, and hypertension; environmental factors like temperature, humidity, and altitude; or even factors affecting individual physical response in the performance of physical activities such as habituation and acclimatization. Recent proposals do not include these factors in perceived exertion assessment, which can lead to inaccurate decision-making in the allocation of a job post, for example.

The proposed method can handle any of these factors; habituation is an example of how this can be done. To handle other factors (like nutrition), variables and their domains have to be known (such as the quantity (kg and liters) and quality (calories) of food ingested and the time (hours and min) spent between the consumption of food and the performance of the activity), as well as their impact on perceived exertion, that is, the weight to be given to each variable in estimating the effort.

Habituation is considered as the training experiences based on tasks repetition, which conduces to a better physical activities' performance and changes in perceived exertion.

It has been stated in recent works that considering individual characteristics like maximum effort in activities' performance conduces to a better perceived exertion assessment. Published results of perceived exertion are based on one of the following methods: Borg method which considers a HR = [60, 220] to define a 14-value scale [22, 28] going from no exertion to maximal exertion, Chamoux method considering person age (220 – age) as the baseline to define a 70-value scale RCC = [0, 69], grouped in sets of 10 values each going from very light to intense, and extended Chamoux which considers individual maximum exertion as the baseline. However, effort level transition in those scales based on HR values may not correspond to perceived

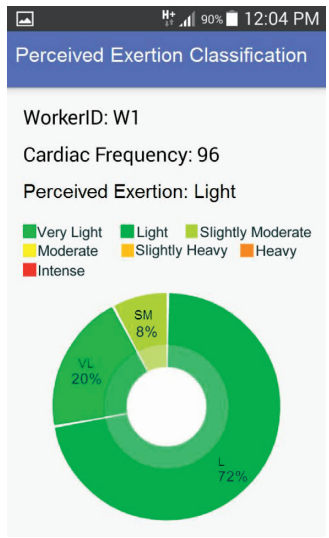


Figure 15: Perceived exertion prototype.

exertion (as perceived by the worker), due to lack of habituation, that is, when considering habituation, a compensation value has to be added to personal exertion for less habituated ones.

Since the assessment of perceived effort and habituation are based on human experience, the fuzzy logic can be used in such evaluations, given that discrete and continuous membership functions of a fuzzy set are intended to capture a person's thinking. Fuzzy membership functions can be determined subjectively in practical problems based on an expert's opinion. Membership functions can be considered as a technique to formalize empirical problem solving that is based on experience rather than the knowledge of theory.

Fuzzy membership function of habituation takes one of three values: no habituated, moderate habituation, and habituated. Fuzzy set values are defined in [0, 1] interval: border values are for a nonhabituated person and a habituated one, respectively. The membership degree was determined by worker experience and direct observation of workers performing physical activities. Values from fuzzy membership function of habituation, CCR, 20% compensation for nonhabituated and 10% compensation for moderate habituated, are inputs for fuzzy membership function of perceived exertion. Fuzzy membership function of perceived exertion takes one of 7 values: very light, light, slightly moderate, moderate, slightly heavy, heavy, and intense, each one taking ten values in [0, 69] interval.

Habituation contribution to perceived exertion assessment is a more realistic result in terms of human experience. As it can be seen in the results (Figure 9), perceived exertion can change to the next level when considering habituation. We think by experience that when a person is not accustomed to perform labor physical activities, the higher the hardness of activities is, the higher the level of perceived effort is. This is true when considering that the membership of one value to a set is binary, that is, a person is habituated or not to perform an activity; however, when considering fuzzy sets, the membership is defined by a function that takes its values in the interval [0, 1]. The closer the degree of membership is to 1, the more the element will be in the set, and the closer the degree of membership is to 0, the less the element will be in the set.

Results obtained in this study clearly establish the importance of considering factors such as habituation to physical activities performance in the evaluation of perceived exertion of workers. Adjustment of perceived exertion levels achieved with fuzzy logic allows us to improve decision-making for the allocation of jobs, the planning of workloads, or even the reduction of risks of fatigue accidents.

8. Conclusions

The contributions of this proposal are the ability to classify perceived exertion of people in daily activities, to improve their safety and health. This is because it is formally established that the effort of a person can be estimated based on his/her cardiac frequency. A standard effort can be estimated for each activity as a reference to analyze the gap with the personalized perceived exertion estimated by our proposed method to perform those activities; the usefulness of measuring the personalized effort of workers in their work environment to preserve their health; and the possibility to determine that a person is conducting his/her activities in a comfortable way, that is, in accordance with his/her personal capacities, abilities, and habituation to improve performance, safety, and welfare state.

This is not a proposal to accurately measure the physical effort but emphasizes the importance of customizing the measurement process and mentions that it is hardly possible to have a generic method, given a large number of variables that must be considered. The intention is to show how the effort estimation varies when considering a custom value as the maximum personal cardiac frequency, as well as imprecision and uncertainty of variables affecting methods to classify perceived exertion.

Using the fuzzy logic, it was possible to verify the importance of the degree of membership of a variable to a fuzzy set, because depending on the degree of membership it is possible that the perceived exertion can be increased (next label) as a result of the rules used by the inference engine. This situation is more suited to real life where although a variable belongs to a certain group, there is a level of belonging to that group, which should be considered because it may be more correct to classify the variable as belonging to a nearby set.

The proposed method for the classification of perceived exertion considers how to add possible variables, as exemplified by the inclusion of habituation and the way it affects, as obtained in related studies. Analysis of our results reveals that an objective method of estimating individual effort should consider custom values in the parameters to capture the widest possible set of variables involved in the estimation of perceived exertion. Therefore, the decision to perform a stress test for obtaining the maximum heart rate is important, because with this action, indirectly, we are including many factors such as age, sex, body mass index, and acclimation.

We conclude that the use of a wearable device with capacities of measurement of physiological parameters together with fuzzy logic computational methods provokes expert knowledge that represent a viable automatic solution for perceived exertion classification.

Future work includes other factors involved such as environment, gender, body mass index, and mental stress. Another type of sensors must be considered, as well as the combination of heterogeneous sensors. The perceived exertion should be objective; direct observation gives an idea of the results, but it is based on experience and questionnaires. Habituation to physical effort requires further study; to our knowledge, there are no studies analyzing the impact of habituation to perform physical activities in the perceived exertion of workers. Additionally, it is important to extend FPC method definition to integrate the implementation into a smartphone.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This paper was supported by Universidad Juárez Autónoma de Tabasco through Grant no. UJAT-2014-IA-01.

References

[1] S. Adibi, *Mobile Health. A Technology Road Map*, Springer International Publishing, Cham, Switzerland, 2015.

[2] R. S. H. Istepanian, C. S. Pattichis, and S. Laxminarayan, *Ubiquitous M-Health Systems and the Convergence Towards 4G Mobile Technologies*, pp. 3–14, Springer, Boston, MA, USA, 2006.

[3] B. J. Noble and R. J. Robertson, *Perceived Exertion*, Human Kinetics Publishers, Champaign, IL, USA, 1996.

[4] A. Holtermann, M. B. Jørgensen, B. Gram et al., "Worksite interventions for preventing physical deterioration among employees in job-groups with high physical work demands: background, design and conceptual model of finale," *BMC Public Health*, vol. 10, no. 1, p. 120, 2010.

[5] S. Gallagher, J. S. Moore, and T. J. Stobbe, *Physical Strength Assessment in Ergonomics*, pp. 1–61, American Industrial Hygiene Association, Falls Church, VA, USA, January 1998.

[6] AIHA Technical Committee, "Ergonomics guide to assessment of metabolic and cardiac costs of physical work," *American Industrial Hygiene Association Journal*, vol. 32, no. 8, pp. 560–564, 1971.

[7] S. Nogareda-Cuixart and P. Luna-Mendaza, *Ntp 323: Determination of Metabolic Rate*, Instituto Nacional de Seguridad e Higiene en el Trabajo. Ministerio de Trabajo y Asuntos Sociales, Espaa, Madrid, Spain, 2005.

[8] K. Parsons, "Occupational health impacts of climate change: current and future iso standards for the assessment of heat stress," *Industrial Health*, vol. 51, no. 1, pp. 86–100, 2013.

[9] S. E. Crouter, J. R. Churilla, and D. R. Bassett, "Estimating energy expenditure using accelerometers," *European Journal of Applied Physiology*, vol. 98, no. 6, pp. 601–612, 2006.

[10] G. A. Borg, "Psychophysical bases of perceived exertion," *Medicine and Science in Sports and Exercise*, vol. 14, no. 5, pp. 377–381, 1982.

[11] D. W. Esliger, A. V. Rowlands, T. L. Hurst, M. Catt, P. Murray, and R. G. Eston, "Validation of the genea accelerometer," *Medicine and Science in Sports and Exercise*, vol. 43, no. 6, pp. 1085–1093, 2011.

[12] P. Pancardo, F. D. Acosta, J. A. Hernández-Nolasco, M. A. Wister, and D. López-de Ipiña, "Real-time personalized monitoring to estimate occupational heat stress in ambient assisted working," *Sensors*, vol. 15, no. 7, pp. 16956–16980, 2015.

[13] P. Pancardo, J. A. Hernández-Nolasco, F. D. Acosta, and M. A. Wister, "Personalizing physical effort estimation in workplaces using a wearable heart rate sensor," in *Proceedings of the 10th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2016)*, Part II 10, pp. 111–122, Springer, San Bartolomé de Tirajana, Gran Canaria, Spain, December 2016.

[14] P. Frimat, M. Amphoux, and A. Chamoux, "Interprétation et mesure de la fréquence cardiaque," *Revue de Medicine du Travail XV*, vol. 147, no. 4, p. 165, 1988.

[15] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 1, pp. 28–45, 1973.

[16] International Labour Organization, "Safety and health at work," 2016, http://www.ilo.org.

[17] ISO Standard, 9886, *Ergonomics-Evaluation of Thermal Strain by Physiological Measurements*, International Standard Organization, Geneva, Switzerland, 2nd edition, 2004.

[18] L. G. Sylvia, E. E. Bernstein, J. L. Hubbard, L. Keating, and E. J. Anderson, "A practical guide to measuring physical activity," *Journal of the Academy of Nutrition and Dietetics*, vol. 114, no. 2, pp. 199–208, 2014.

[19] S. J. Strath, A. M. Swartz, D. R. Bassett Jr., W. L. O'Brien, G. A. King, and B. E. Ainsworth, "Evaluation of heart rate as a method for assessing moderate intensity physical activity," *Medicine and Science in Sports and Exercise*, vol. 32, no. 9, pp. S465–S470, 2000.

[20] H. Medjahed, D. Istrate, J. Boudy et al., "A fuzzy logic approach for remote healthcare monitoring by learning and recognizing human activities of daily living," *Fuzzy Logic-Emerging Technologies and Applications*, InTech, London, UK, 2012.

[21] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, "Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence," *IEEE Transactions on Automatic Control*, vol. 42, no. 10, pp. 1482–1484, 1997.

[22] M. E. Bouton, *Learning and Behavior: A Contemporary Synthesis*, Sinauer Associates, Sunderland, MA, USA, 2007.

[23] J. Borresen and M. I. Lambert, "The quantification of training load, the training response and the effect on performance," *Sports Medicine*, vol. 39, no. 9, pp. 779–795, 2009.

[24] E. M. Heath, J. R. Blackwell, U. C. Baker, D. R. Smith, and K. W. Kornatz, "Backward walking practice decreases oxygen uptake, heart rate and ratings of perceived exertion," *Physical Therapy in Sport*, vol. 2, no. 4, pp. 171–177, 2001.

[25] J. F. Patton, W. P. Morgan, and J. A. Vogel, "Perceived exertion of absolute work during a military physical training program," *European Journal of Applied Physiology and Occupational Physiology*, vol. 36, no. 2, pp. 107–114, 1977.

[26] S. Hwang, J. Seo, J. Ryu, and S. Lee, "Challenges and opportunities of understanding construction worker's physical demands through field energy expenditure measurements using a wearable activity tracker," in *Proceedings of Construction Research Congress 2016*, pp. 2730–2739, San Juan, Puerto Rico, May–June 2016.

[27] A. Kolus, D. Imbeau, P.-A. Dubé, and D. Dubeau, "Adaptive neuro-fuzzy inference systems with k-fold cross-validation for energy expenditure predictions based on heart rate," *Applied Ergonomics*, vol. 50, pp. 68–78, 2015.

[28] G. C. Migliaccio, J. Teizer, T. Cheng, and U. C. Gatti, "Automatic identification of unsafe bending behavior of construction workers using real-time location sensing and physiological status monitoring," in *Proceedings of the Construction Research Congress*, vol. 2123, p. 633642, West Lafayette, IN, USA, May 2012.

[29] A. Aryal, A. Ghahramani, and B. Becerik-Gerber, "Monitoring fatigue in construction workers using physiological measurements," *Automation in Construction*, vol. 82, pp. 154–165, 2017.

[30] Y. Shimizu and T. Jindo, "A fuzzy logic analysis method for evaluating human sensitivities," *International Journal of Industrial Ergonomics*, vol. 15, no. 1, pp. 39–47, 1995.

[31] W. Karwowski and M. M. Ayoub, "Fuzzy modelling of stresses in manual lifting tasks," *Ergonomics*, vol. 27, no. 6, pp. 641–649, 1984.

[32] E. Tóth-Laufer, M. Takács, and I. J. Rudas, "Fuzzy logic-based risk assessment framework to evaluate physiological parameters," *Acta Polytechnica Hungarica*, vol. 12, no. 2, pp. 159–178, 2015.

[33] R. M. T. Laukkanen and P. K. Virtanen, "Heart rate monitors: state of the art," *Journal of Sports Sciences*, vol. 16, no. 1, pp. 3–7, 1998.

Hindawi
Mobile Information Systems
Volume 2018, Article ID 7381264, 10 pages
https://doi.org/10.1155/2018/7381264

Research Article

Gait Analysis Using Computer Vision Based on Cloud Platform and Mobile Device

Mario Nieto-Hidalgo, Francisco Javier Ferrández-Pastor, Rafael J. Valdivieso-Sarabia, Jerónimo Mora-Pascual, and Juan Manuel García-Chamizo

Department of Computing Technology, University of Alicante, Campus San Vicente del Raspeig, Alicante, Spain

Correspondence should be addressed to Mario Nieto-Hidalgo; mnieto@dtic.ua.es

Received 30 June 2017; Revised 31 October 2017; Accepted 12 November 2017; Published 14 January 2018

Academic Editor: Pino Caballero-Gil

Frailty and senility are syndromes that affect elderly people. The ageing process involves a decay of cognitive and motor functions which often produce an impact on the quality of life of elderly people. Some studies have linked this deterioration of cognitive and motor function to gait patterns. Thus, gait analysis can be a powerful tool to assess frailty and senility syndromes. In this paper, we propose a vision-based gait analysis approach performed on a smartphone with cloud computing assistance. Gait sequences recorded by a smartphone camera are processed by the smartphone itself to obtain spatiotemporal features. These features are uploaded onto the cloud in order to analyse and compare them to a stored database to render a diagnostic. The feature extraction method presented can work with both frontal and sagittal gait sequences although the sagittal view provides a better classification since an accuracy of 95% can be obtained.

1. Introduction

This work is part of a project called Gait-A whose main objective is the early detection of frailty and senility syndromes using gait analysis. Physical activity is one of the main components involved in frailty syndrome evaluation [1, 2]. Gait is identified as a high cognitive task in which attention, planning, memory, and other cognitive processes are involved [3, 4].

Through gait analysis, quantification of measurable information of gait, and its interpretation [5], frailty and dementia syndromes can be diagnosed. This process is carried out by specialists and is based on estimations through visual inspection of gait.

In this work, we propose a computer vision approach that could aid the specialists providing them with objective measurements of gait and, thus, gain in objectivity of the gait analyses performed.

We propose the use of smartphone cameras to record the subject's gait and also provide computer vision algorithms able to analyse those sequences to extract spatiotemporal gait parameters. These parameters are then sent to the cloud to be analysed by a classifier for the purpose of determining whether abnormalities are present or not.

A lot of works dealing with gait analysis using computer vision are found in the literature. However, most of them focus on gait biometrics for human identification, and few of them address gait analysis for detection of abnormalities.

The main goal of this study is to provide a nonexpensive and easy-to-deploy solution to obtain the spatiotemporal parameters of gait, which will be fed to classification algorithms that will discriminate between normal and abnormal gait. It needs to be mentioned that the process of obtaining spatiotemporal parameters for abnormal gait compounds the task as the number of assumptions that can be made over gait patterns is drastically reduced. In such cases, neither cyclic patterns nor the totality of the gait phases can be assumed to be present. In this work, for study purposes, Parkinsonian gait, knee pain, and foot dragging among other patterns that deviate from what we consider normal gait will be taken as abnormal gait.

A set of different gait features is analysed in [6] for person identification. The process starts by extracting the silhouette with a background subtraction technique to then obtain the contour. After the contour is obtained, they extract four time-series features: width/height ratio, bounding box width, silhouette area, and center of gravity (COG). These four features

follow a cyclic pattern that match the gait cycle and are used to identify a person through deterministic learning.

Xu et al. examined the suitability of the Kinect sensor to measure gait parameters while walking on a treadmill in frontal view [7]. They compared the heel strike (HS) and toe off (TO) they obtained with those obtained using a motion tracking system. HS showed less error than TO because it happens closer to the sensor.

Choudhury and Tjahjadi [8] proposed a method composed of three modules: silhouette extraction, subject classification using Procrustes shape analysis (PSA) and elliptic Fourier descriptor (EFD), and combination of both results. For silhouette extraction, they use background subtraction and morphologic operations to remove noise. PSA module analyses a group of shapes using matching of geometrical locations of a silhouette. The stride length is computed using the width of the bounding box. Finally, EFD allows to characterize the contour of the subject in key points of a gait phase.

Leu et al. proposed a method to extract skeleton joints from sagittal and frontal views [9]. The method proposed uses the horizontal and vertical projection of the silhouette pixels to obtain the neck joint. Then they apply an anatomical model to obtain hip, knees, and ankles. Yoo and Nixon [10] also extract skeleton joints using an anatomical model to segment the silhouette but they obtain the mean points of each segment and then apply linear regression to obtain a line that represents the bones. During double support gait phase, they apply motion tracking to estimate the location of the occluded points. Khan et al. [11], similarly obtain the skeleton by computing the mean points of each body segment. They obtain leg movement and posture inclination and compare it with a normal gait model to recognise Parkinsonian gait.

In addition, we find the following proposals for classifying gait patterns. In Wang [12], the method is based on optical flow that calculates a histogram of silhouette flows to which an eigenspace transformation applies. The data obtained are compared with a normal gait template to calculate deviation. In Bauckhage et al. [13], homeomorphisms apply between 2D lattices and binary shapes to obtain a vector space in which the silhouette is encoded. They performed several silhouette bounding box splittings to obtain different lattices that are then classified using support vector machine (SVM).

Apparently, most of the vision-based gait analysis proposals use sagittal view for the reason that it provides more information with which to work. However, there are obtainable benefits out of a frontal gait analysis. According to Whittle [14], more gait abnormalities can be observed from a sagittal view than from a frontal view. However, we do also undertake frontal gait analysis for the following reasons:

(i) Some abnormalities can only be observed from a frontal point of view. Whittle [14] mentions that circumduction gait, hip hiking, abnormal foot contact, and rotation among others are better observed from a frontal view.

(ii) In terms of the physical space necessary for recording, sagittal gait sequences require much more than those of frontal gait, for which only a small hall or corridor will serve.

A way to reduce the space needed for sagittal view recording is to use a treadmill, but it could alter gait patterns, especially with frail people. Another workaround is to use a motorised camera that follows the subject, but it is expensive and could complicate the background subtraction as it is moving as well. Both workarounds complicate the acquisition of gait sequences making it difficult to be processed by a smartphone.

Sagittal images show a clear view of feet displacement and enough information to locate heel and toe of each foot. In frontal view, on the other hand, it is not easy to determine where the heel and toe are located in each foot. Therefore, a different approach is required for frontal sequences.

In sagittal view, the size of the subject's silhouette is maintained along the whole of its trajectory. However, in frontal view, the size of the silhouette increases along its trajectory, so a normalization might be required.

The paper is organized as follows. Section 2 describes the sagittal and frontal methods to obtain spatiotemporal parameters of gait, their implementation in a smartphone, and the classification of normal and abnormal gait in a cloud platform. Section 3 shows the results in which the spatiotemporal gait parameters are subjected to normal and abnormal gait classification. Finally, Section 4 provides the conclusion of this work.

2. Methods

In this paper, we present a platform for gait analysis using computer vision where a smartphone records and processes a gait sequence to obtain spatiotemporal parameters to be sent to the cloud for a classification between normal and abnormal gait. The layout of the platform is shown in Figure 1. In the following subsections, each module of the platform will be described.

2.1. Sagittal Approach. The sagittal approach takes gait sequences recorded from the side as input. The method presents four phases: preprocessing, feet location, feature extraction, and skeleton extraction. Figure 2 shows the diagram of the sagittal approach. The classification phase is performed in the cloud.

2.1.1. Preprocessing. In this phase, a background subtraction is performed to obtain the silhouette of the subject using mixture of Gaussians [15] background subtraction. After that, a morphology operator is applied to remove noise. Finally, the bounding box of the remaining silhouette is extracted by computing the x, y positions using (1), and then those points are made to correspond to a rectangle $(x, y, \text{width}, \text{height})$ using (2).

$$\begin{aligned}
\min x &= \arg\min_{x,y} \left(\forall_{x,y} \in \text{silhouette} : x\right) \\
\max x &= \arg\max_{x,y} \left(\forall_{x,y} \in \text{silhouette} : x\right) \\
\min y &= \arg\min_{x,y} \left(\forall_{x,y} \in \text{silhouette} : y\right) \\
\max y &= \arg\max_{x,y} \left(\forall_{x,y} \in \text{silhouette} : y\right),
\end{aligned} \tag{1}$$

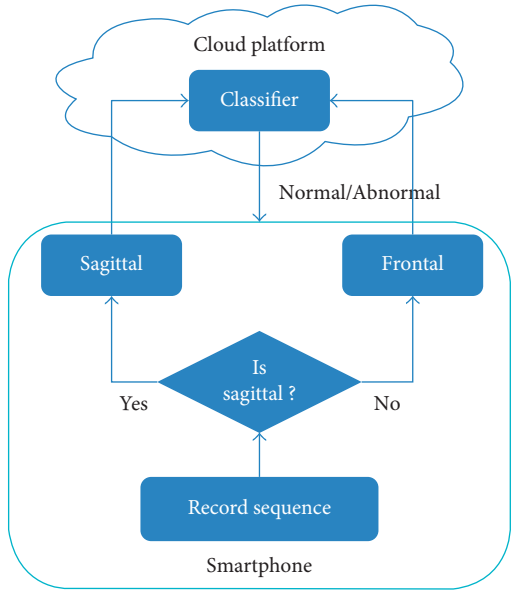


Figure 1: Diagram of the proposed platform. Extraction of gait features is performed on a smartphone and classification in the cloud.

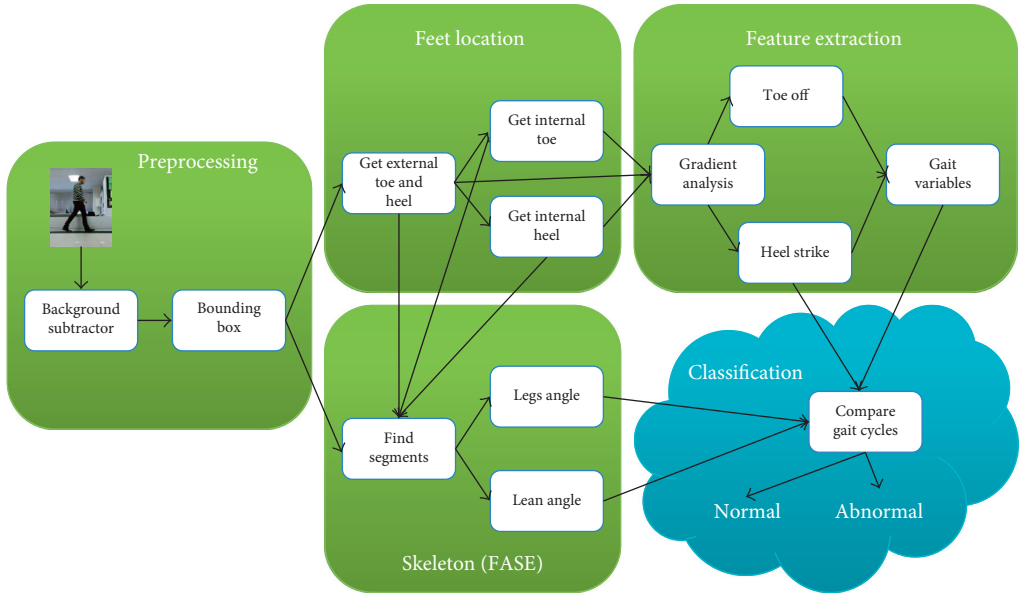


Figure 2: Diagram of the sagittal gait approach.

$$\text{Bounding box} = (\min x, \min y, \max x - \min x, \max y - \min y). \tag{2}$$

2.1.2. Feet Location. The silhouette obtained by background subtraction is then enclosed in its bounding box and split into four regions, namely, *head (13% of bounding box height), torso (34%), upper legs (24%), and lower legs (29%)*, according to an anthropometric model [16] as shown in Figure 3. The lower leg region is then brought to focus. We search the silhouette pixel with maximum X component to obtain the toe of the front foot (FF) using (3) and the pixel with minimum X to obtain the heel of the back foot (BF) using (4). Then, the lower leg region is split into halves vertically to separate each foot. In the BF half, we search for the lower right pixel (assuming displacement from left to right) to obtain the BF toe. In the FF half, we search for the lower left pixel to obtain the heel. The final result is shown in Figure 3.

$$\arg\max_{x,y} \left(\forall_{x,y} \in \text{silhouette} : x\right), \tag{3}$$

$$\arg\min_{x,y} \left(\forall_{x,y} \in \text{silhouette} : x\right). \tag{4}$$

2.1.3. Feature Extraction. For each frame of the sequence, the position of the heel and toe of both feet was obtained in

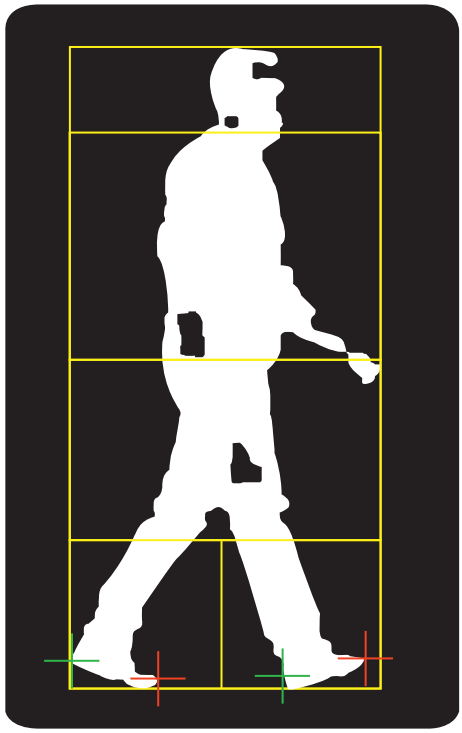


Figure 3: Location of the heel and toe of each foot for the sagittal approach.

the previous phase. To these time series, we applied gradient analysis of the X component to obtain heel strike (HS) when the mean point gradient between FF heel and FF toe goes from greater than zero to zero (foot stops moving as shown in (5)) and the toe off (TO) when the mean point gradient between BF heel and BF toe goes from zero to greater than zero (foot starts moving as shown in (6)). Applying the gradient directly over the position time series produces a lot of false positives due to some noise. To filter the noise, we apply a threshold where any gradient value less than that is set to zero. This threshold can remove small oscillations due to an error in the process of getting the silhouette and locating toes and heels. It follows that a Gaussian smoothing is applied, and isolated values greater than zero or equal to zero are removed using (7).

$$\text{heel\_strike}(i) \rightarrow \text{grad}(x)_i \leq 0 \wedge \text{grad}(x)_{i-1} > 0, \quad (5)$$

$$\text{toe\_off}(i) \rightarrow \text{grad}(x)_i > 0 \wedge \text{grad}(x)_{i-1} = 0, \quad (6)$$

$$v_i' = \begin{cases} \text{if } v_i = 0 \wedge v_{i-1} \neq 0 \wedge v_{i+1} \neq 0 \rightarrow \dfrac{(v_{i-1} + v_{i+1})}{2} \\ \\ \text{if } v_i \neq 0 \wedge v_{i-1} = 0 \wedge v_{i+1} = 0 \rightarrow 0. \end{cases} \quad (7)$$

2.1.4. Skeleton Extraction. The skeleton extraction phase provides a fast way of obtaining an approximation of the locations of the head, neck, hip, knees, and feet. It uses the four regions of the silhouette described in the feet location phase. The head and torso regions are divided in half horizontally, and the COG of each half is computed. The COG of the upper region is moved to the top, and the COG of the lower region is moved to the bottom. Then, the head lower COG and the torso upper COG are averaged to obtain a common point which is the neck. The head location corresponds to the upper COG of the head region.

The upper leg region is also split horizontally in half, and both COGs are obtained. In addition, a vertical split is also performed, and another two COGs are obtained. The upper COG is moved to top and averaged with the lower torso COG to obtain the hip location. Lower COG is discarded. Then right and left COGs are moved to bottom, those two points being the location of the knees. The knees are adjusted to simulate bending. The process to adjust the knees consists in tracing three circles: one with center at the hip and thigh length radius (which is the height of the upper leg segment) and two other circles with center at each foot and radius equal to the tibia length (which is the height of the lower leg segment). Then, an intersection between the hip circle and each of the foot circles is performed. There are three possibilities:

(i) *No intersection*. In this case, the knee point is the one given by the COG.

(ii) *One intersection*. In this case, the knee point is the intersection point.

(iii) *Two intersections*. In this case, the knee point is the intersection point more to the right (assuming gait direction from left to right).

Finally, the location of each foot is the mean point of the heel and toe obtained in the feet location phase. Figure 4 shows the final result.

2.2. Frontal Approach. The frontal approach is very similar to the sagittal one proposed in the previous subsection. It has the same phases: preprocessing, feet location, feature extraction, and skeleton detection. The diagram of the frontal gait approach is shown in Figure 5.

2.2.1. Preprocessing. This phase is exactly the same as for sagittal. The silhouette is obtained using Mixture of Gaussians as background subtraction, and then morphology operators are applied to remove noise.

2.2.2. Feet Location. In frontal view, both toes are always visible but heels are constantly occluded, so heels cannot be properly located. Therefore, we can only rely on toe information.

To obtain toes, we proceed by dividing the silhouette in four regions according to the anthropometric model established in [16]. We focus only on the lower leg segment. Then, we calculate its bounding box and split it vertically into half to separate both feet. It is important to recalculate the bounding box of this part so the vertical split separates both feet accurately; otherwise, any misalignment can cause problems. Note that the

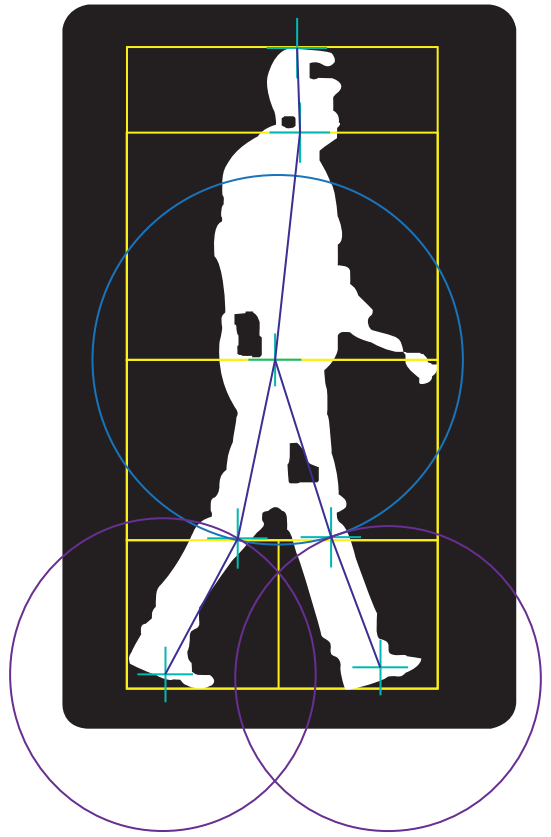


Figure 4: Knee adjustment for the sagittal approach.

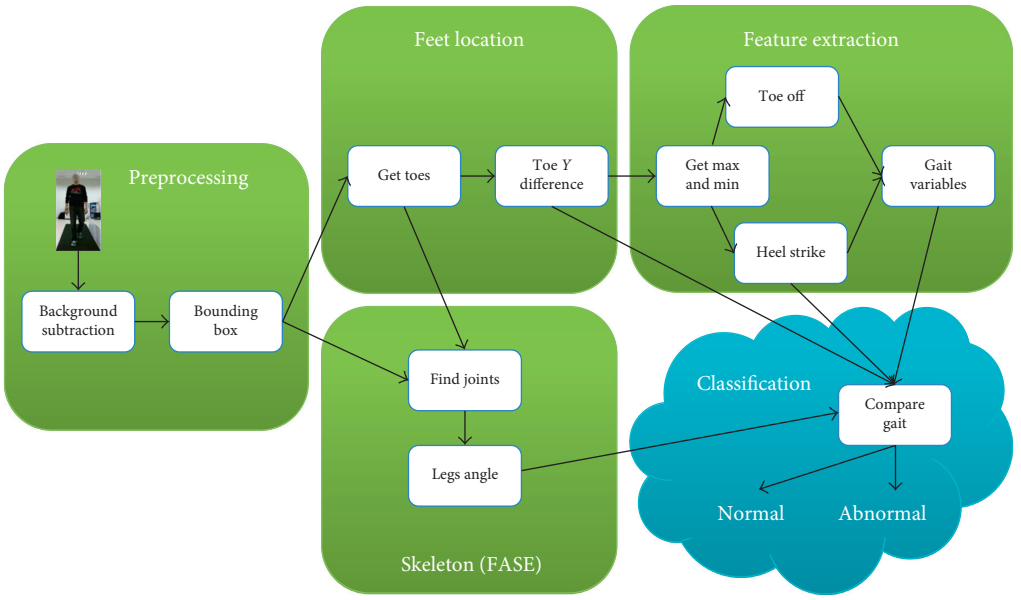


Figure 5: Diagram of the frontal gait approach.

process of splitting the bounding box for the purpose of separating both feet will never be accurate with gait patterns that place one foot in front of the other. We will assume that this particular gait pattern is not present in our dataset. We obtain the left and right foot toe by locating the pixel with minimum y component in the left and right half, respectively (8) (Figure 6).

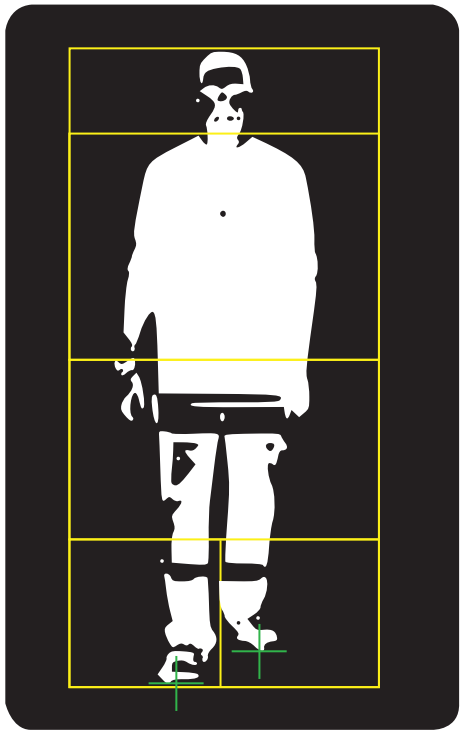


Figure 6: Toe location of each foot for the frontal approach.

$$\arg\min_{x,y}\left(\forall_{x,y} \in \text{silhouette} : y\right). \tag{8}$$

2.2.3. Feature Extraction. The previous phase provides the position of each toe for each frame, which is precisely the information we need to derive HS and TO. We propose an approach to obtain HS and TO with frontal gait based on the time series derived by subtracting the vertical component of both feet.

We will use the subtraction of the y component of the toes to obtain a curve in which zero crosses indicate the feet adjacent gait phase. HS and TO of each foot are located between each zero cross. We can estimate HS and TO by assuming that HS is produced before TO; HS is produced in the first half of each region and TO in the second half. Therefore, we can estimate HS and TO following (9) and (10), respectively, where zc_i relates to the frame in which a zero cross point occurs and zc_{i-1} relates to the frame of the previous zero cross point.

$$\text{HS} = zc_{i-1} + \frac{\left(zc_i - zc_{i-1}\right) \times 3}{4}, \tag{9}$$

$$\text{TO} = zc_i - \frac{zc_i - zc_{i-1}}{4}. \tag{10}$$

This approach poses some problems with some abnormal gait patterns, as shown in [17], in which some events could not be detected, for example, when a foot is always behind the other or is dragged due to some injury or pain. Figure 7 shows foot dragging where, in some cases, the curve does not cross zero during the swing phase. To solve the problem, we devise another method. Using the same curve from the previous approach (the difference of y component of each foot), we proceed by applying Gauss filters to remove noise (Figure 8 shows the curve of Figure 7 after applying Gauss filters), then we obtain the local maxima and minima, which are located more or less at the center of each pair of zero crosses. But, in this case, the curve does not have to cross zero to produce a maximum or minimum, and the problem is solved.

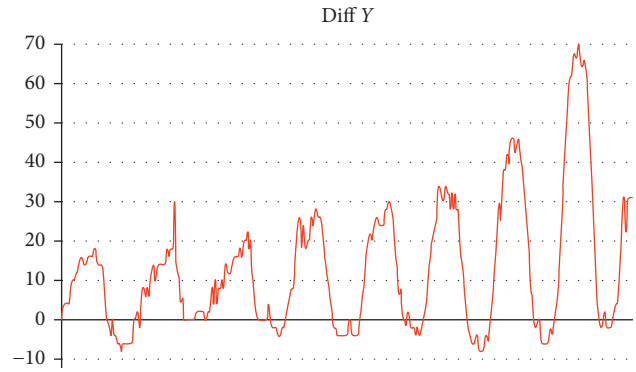


Figure 7: Difference of component Y of each foot with abnormal foot dragging.

HS are located before a maximum or minimum, and TO after. We know that both events are located in that region. Empirically adjusting them, we derived that the HS is located at 1/4 the distance between one maximum (or minimum) and the previous one (12), and TO is located at 1/8 the distance between one maximum (or minimum) and the next one (13).

Being M an ordered set of maxima and minima in ascending chronological order:

$$M = \{m_1, m_2, m_3 \cdots m_n\}. \tag{11}$$

HS of m_i is obtained as

$$\text{HS}_i = m_i - \frac{m_i - m_{i-1}}{4}, \tag{12}$$

and TO is obtained as

$$\text{TO}_i = m_i + \frac{m_{i+1} - m_i}{8}. \tag{13}$$

2.2.4. Skeleton Detection. The process is the same as the one described for the sagittal approach, but for frontal approach, the adjustment of knees is not necessary.

2.3. Smartphone Implementation. Sagittal and frontal approaches were implemented on Android using OpenCV native functions. We allowed two ways of processing a dataset:

(i) On a real-time video: the smartphone camera records the subject walking and processes it at the same time.

(ii) On a previously recorded video: the smartphone records the subject walking and stores it in memory, and then the stored video is processed.

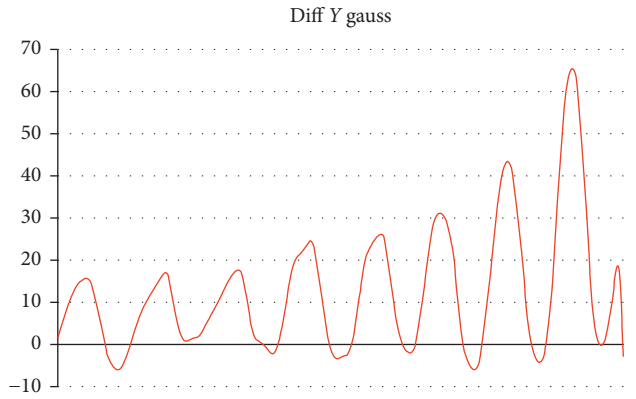


Figure 8: Difference of component Y of each foot after applying Gauss filters.

To achieve real-time processing, we use the pyramidal multirresolution approach described in [18]. We achieve 10 fps using a quad core at 1.4 GHz smartphone with 1 GB memory and 25 fps using a tablet with a Tegra K1 quad core processor at 2.2 GHz and 2 GB memory. The size of the input image was reduced to 480×270 pixels. However, results shown in Section 3 are obtained using full resolution using the dataset.

2.4. Cloud Platform. To develop the cloud platform, we used the Microsoft Azure Machine Learning platform. This is a cloud platform for designing and developing predictive models. Azure provides a REST Web Service to access the Machine Learning tools.

For our purposes, we develop a K-nearest neighbour (KNN) algorithm with Dynamic Time Warping (DTW) as a distance function accessed through the REST Web Service provided by Azure. To perform a classification between normal and abnormal gait, we use the stride (bounding box width for sagittal approach, and subtraction between y component of each foot for frontal approach) and leg-angle time series (provided by the skeleton extraction algorithm computed as the angle formed by the hip and each foot).

3. Results and Discussion

We will now describe the experiments performed and the results obtained. The dataset recorded for the experiments is also described in this section.

3.1. Dataset. To test the proposed approaches, we recorded two datasets of subjects walking: one using sagittal view and the other using frontal view. Both datasets were recorded in a room with a nonhomogeneous background including windows where the light made it difficult to extract the silhouette. This was intentional because we wanted to test our approaches in real conditions, and so the silhouette is often incomplete. Figure 9 shows the room in which the recordings were performed.

To record the frontal dataset, we placed a camera at one end of an 8 m corridor and asked the subject to walk towards it.

Figure 9: Dataset background showing a room with closets and windows.

We captured a total of 23 samples of normal gait and 20 samples of abnormal. To record the sagittal dataset, we used the same environment, but we placed a camera at a distance of 4 m from the perpendicular of the gait direction to obtain a side view. In this case, a total of 15 samples of normal gait and 15 of abnormal gait were recorded. Even if the number of recorded samples is low (43 for frontal gait and 30 for sagittal gait), there are a total of 320 HS events and 319 TO events for frontal gait and 233 HS events and 223 TO events for sagittal gait.

We asked the subjects to walk normally along the corridor and then to walk feigning some of the following abnormalities:

(i) *Knee pain*: the subject simulated pain in one of his knees.

(ii) *Foot dragging*: the subject dragged one foot.

(iii) *Parkinsonian gait*: the subject made some small steps with variable speed.

(iv) *Other*: the subject depicted random patterns.

To guarantee the privacy of the subjects, we published only the silhouettes extracted during the silhouette extraction phase. These silhouettes are stored as an ordered set of images, and a file with the elapsed milliseconds for each image is also included. For each recorded sample, we manually mark the frames in which a HS or TO event occurs to use it as a ground truth. We also include information related to pixel width to be able to calculate distances and the sample class (normal = 0 or abnormal = 1). In addition, a file with the output of the feet location and feature extraction phases is included which contains the positions of heel and toe of each foot, their gradients, and the events of HS and TO detected. These results are the output of the HS and TO detection algorithm using full resolution (1920×1080), which do not correspond to those provided by the smartphone using a quarter of that resolution.

Both datasets are accessible through the URL provided by [19].

3.2. Experiments. We performed experiments using our own datasets for sagittal and frontal gait. We used the manual marking of the HS and TO events of each gait sequences of the dataset as ground truth. The error margin of this manual marking was set to ± 1 frame because that is the minimum value. We also assumed an error of ± 1 frame in the algorithm

Table 1: Results of the sagittal HS and TO detection algorithm showing the amount of correct detections (less than 2 frames of difference between algorithm and manual marking), undetected cases, wrong detection (more than 2 frames of difference), and the root mean square error of both correct and wrong cases.

Approach	Correct	Undetected	Wrong	RMSE
DAI dataset normal gait heel strike				
Sagittal	90.2%	1.1%	7.6%	1.44 frames (48 ms)
Frontal	89.4%	0%	10.60%	1.88 frames (63 ms)
Toe off				
Sagittal	93.3%	2.2%	2.2%	1.08 frames (36 ms)
Frontal	89.4%	0%	10.6%	1.63 frames (54 ms)
DAI dataset abnormal gait heel strike				
Sagittal	89%	2.1%	6.9%	1.79 frames (60 ms)
Frontal	72.1%	0%	27.9%	2.42 frames (81 ms)
Toe off				
Sagittal	82.1%	3.6%	10.7%	1.59 frames (53 ms)
Frontal	75%	0%	25%	2.17 frames (72 ms)
Total heel strike				
Sagittal	89.5%	1.7%	7.2%	1.66 frames (55 ms)
Frontal	78.8%	0%	21.3%	2.23 frames (74 ms)
Toe off				
Sagittal	86.5%	3.0%	7.4%	1.41 frames (47 ms)
Frontal	80.6%	0%	19.4%	1.98 frames (66 ms)

output. So, the global error margin was set to ± 2 frames. Then, the difference in frames between the ground truth and the proposed algorithm was analysed. Any difference less or equal to the global error margin was considered acceptable. Then, the root mean square error (RMSE) of the differences was computed using

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(m_i - a_i)^2}, \tag{14}$$

where n corresponds to the number of events (HS or TO in this case), m_i the frame of the event i in the manual marking, and a_i the frame of the event i in the algorithm output.

3.3. Sagittal Approach. In Table 1, we show the results after applying the HS and TO detection algorithm with the filtering method described in the previous section for sagittal view. The table shows the amount of correct detections (less than 2 frames of difference between algorithm and manual marking), undetected cases, wrong detection (more than 2 frames of difference), and the root mean square error of both correct and wrong cases. As observed, the RMSE of both HS and TO events is lower than the error margin of 2 frames. TO events are more accurately delimited than HS events. But, HS events show less undetected cases. Therefore, it will be HS, the event we will use to obtain the spatiotemporal parameters to perform classification. Figure 10 shows graphically the correct, wrong, and undetected cases.

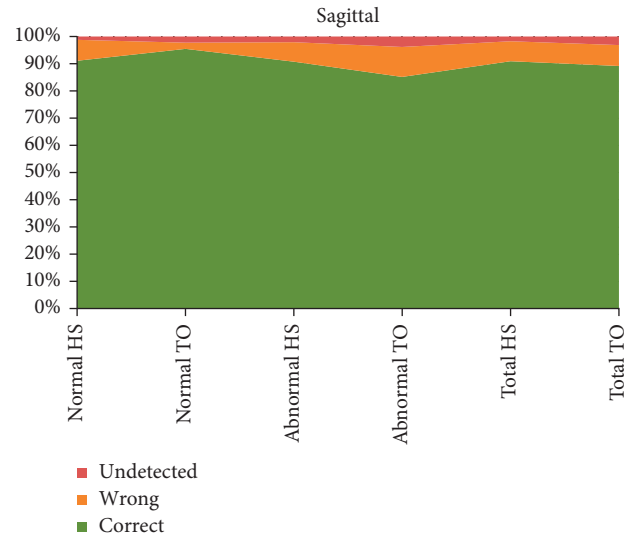


Figure 10: Results of the sagittal approach.

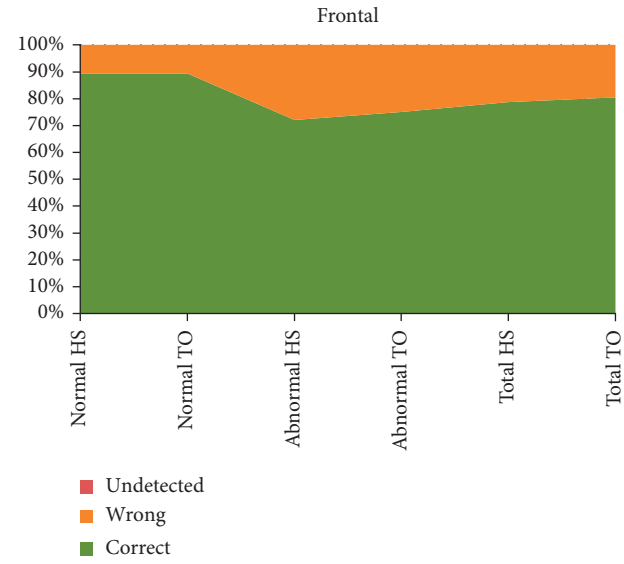


Figure 11: Results of the frontal approach.

3.4. Frontal Approach. Table 1 also shows the results after applying the frontal approach. As shown in there, the RMSE of both HS and TO in normal gait is smaller than the error margin of 2 frames, but it is slightly bigger for abnormal gait. Therefore, results are acceptable for both normal and abnormal. Error is mainly produced in the first steps when the silhouette is smaller (the subject is farthest from the camera). Figure 11 shows graphically the results of Table 1.

The results obtained with our sagittal view approach are similar for normal gait. We obtained 1.44 frames for HS and 1.08 for TO, which were slightly more precise than the ones we extracted from frontal approach (1.88–1.63) but close to each other. However, in the case of abnormal gait, we obtained 1.79 frames for HS and 1.59 for TO, which were more precise than those obtained with the frontal approach (2.42–2.17).

Table 2: 10-fold and leave-one-out cross-validation results of each classifier for the sagittal approach.

Classifier	10-Fold	Leave-one-out
Stride each cycle	77%	75%
Legs angle each cycle	92%	92%
Stride each subject	77%	77%
Legs angle each subject	100%	100%

Table 3: 10-fold and leave-one-out cross-validation results of each classifier for the frontal approach.

Classifier	10-Fold	Leave-one-out
Stride each cycle	75%	79%
Legs angle each cycle	71%	77%
Stride each subject	80%	84%
Legs angle each subject	88%	88%

3.5. Classification. To perform a classification between normal and abnormal gait, we use KNN to compare the stride length and leg-angle time series of the different gait cycles. To calculate the distance between two time series, we apply DTW. We perform the classification test with two different methods:

(i) Testing each gait cycle separately. The time series corresponding to each gait cycle is treated separately as if it belonged to different subjects.

(ii) Testing each gait cycle of each recording sample and outputting the mode class for each subject. In this case, a prediction for each gait cycle follows, and then another prediction is computed by outputting the mode class for the same recording sample.

To validate the proposed classification, we use 10-fold and leave-one-out cross-validations to finely measure the accuracy of each classifier.

Table 2 shows the results of the stride and leg-angle time series for the sagittal approach. We obtained an accuracy rate of 100% using leg-angle time series when outputting the mode class for each recording sample. Least accurate results, however, are the ones offered by the stride width.

The results of the classification experiments for frontal approach are shown in Table 3. As shown in there, testing each recording sample produces better results as it tends to eliminate outliers.

We have focussed on obtaining a classification between normal and abnormal gait to assess the suitability of the proposed algorithm to differentiate between the two of them. For this test, we considered knee pain and foot dragging as abnormal gait. The results obtained suggest that the classifier can differentiate between normal and abnormal gait. Therefore, future work will focus on classifying different abnormal gaits.

4. Conclusion

The main contribution of this paper is a nonexpensive and easy-to-deploy approach to obtain HS and TO and some skeleton joints using both sagittal and frontal gait sequences. Frontal view poses some problems when obtaining heels position, so we focus on toes instead. Results show acceptable precision in providing HS and TO in both the sagittal and the frontal methods. Comparing both approaches, results were similar but sagittal proved to be more accurate. The dataset recorded to test the proposed approaches is for anyone to use it [19]. To maintain the privacy of the subjects, we published only the silhouette.

We also provide a cloud platform-based web service to perform a classification between normal and abnormal gait for both sagittal and frontal views. Results show a classification rate greater than 80% in frontal view and more than 90% in sagittal view.

The ability to perform gait analysis using frontal view reduces the physical space required for the tests. In addition, this method does not rely on silhouette displacement (the sagittal approach does), so it is also suitable for treadmill gait sequences. Therefore, the space could be reduced even more in cases where the alteration of gait patterns that the treadmill could cause does not significantly matter.

Future work will focus on improving the accuracy of HS and TO for abnormal gait and classifying different abnormal gait types.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research is part of the FRASE MINECO project (TIN2013-47152-C3-2-R) funded by the Ministry of Economy and Competitiveness of Spain.

References

[1] J. Waltson and L. Fried, "Frailty and the old man," *Medical Clinics of North America*, vol. 83, no. 5, pp. 1173–1194, 1999.

[2] L. P. Fried, C. M. Tangen, J. Walston et al., "Frailty in older adults evidence for a phenotype," *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 56, no. 3, pp. M146–M157, 2001.

[3] J. M. Hausdorff, G. Yogev, S. Springer, E. S. Simon, and N. Giladi, "Walking is more like catching than tapping: gait in the elderly as a complex cognitive task," *Experimental Brain Research*, vol. 164, no. 4, pp. 541–548, 2005.

[4] T. Mulder, W. Zijlstra, and A. Geurts, "Assessment of motor recovery and decline," *Gait & Posture*, vol. 16, no. 2, pp. 198–210, 2002.

[5] F. I. Mahoney and D. W. Barthel, "Functional evaluation: the barthel index," *Maryland State Medical Journal*, vol. 14, pp. 61–65, 1965.

[6] W. Zeng, C. Wang, and F. Yang, "Silhouette-based gait recognition via deterministic learning," *Pattern Recognition*, vol. 47, no. 11, pp. 3568–3584, 2014.

[7] X. Xu, R. W. McGorry, L. S. Chou, J. H. Lin, and C. C. Chang, "Accuracy of the microsoft kinect™ for measuring gait parameters during treadmill walking," *Gait & Posture*, vol. 42, no. 2, pp. 145–151, 2015.

[8] S. D. Choudhury and T. Tjahjadi, "Gait recognition based on shape and motion analysis of silhouette contours," *Computer Vision and Image Understanding*, vol. 117, no. 12, pp. 1770–1785, 2013.

[9] A. Leu, D. Ristić-Durrant, and A. Gräser, "A robust markerless vision-based human gait analysis system," in *Proceedings of the 6th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI) 2011*, pp. 415–420, IEEE, Timişoara, Romania, May 2011.

[10] J. H. Yoo and M. S Nixon, "Automated markerless analysis of human gait motion for recognition and classification," *ETRI Journal*, vol. 33, no. 2, pp. 259–266, 2011.

[11] T. Khan, J. Westin, and M. Dougherty, "Motion cue analysis for parkinsonian gait recognition," *The Open Biomedical Engineering Journal*, vol. 7, no. 1, pp. 1–8, 2013.

[12] L. Wang, "Abnormal walking gait analysis using silhouette-masked flow histograms," in *Proceedings of the 18th International Conference on Pattern Recognition, ICPR 2006*, vol. 3, pp. 473–476, IEEE, Cancun, Yucatán, Mexico, August 2006.

[13] C. Bauckhage, J. K. Tsotsos, and F. E. Bunn, "Automatic detection of abnormal gait," *Image and Vision Computing*, vol. 27, no. 1-2, pp. 108–115, 2009.

[14] M. W. Whittle, *Gait Analysis: An Introduction*, Butterworth-Heinemann, Oxford, UK, 2014.

[15] Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recognition Letters*, vol. 27, no. 7, pp. 773–780, 2006.

[16] F. Tafazzoli and R. Safabakhsh, "Model-based human gait recognition using leg and arm movements," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 8, pp. 1237–1246, 2010.

[17] M. Nieto-Hidalgo, F. J. Ferrández-Pastor, R. J. Valdivieso-Sarabia, J. Mora-Pascual, and J. M. García-Chamizo, "Vision based gait analysis for frontal view gait sequences using rgb camera," in *Proceedings of the Ubiquitous Computing and Ambient Intelligence: 10th International Conference, UCAmI 2016, Part I*, vol. 10, pp. 26–37, Springer, San Bartolomé de Tirajana, Gran Canaria, Spain, November–December 2016.

[18] M. Nieto-Hidalgo and J. M. García-Chamizo, "Real time gait analysis using rgb camera," in *Proceedings of the Ubiquitous Computing and Ambient Intelligence: 10th International Conference, UCAmI 2016, Part I*, vol. 10, pp. 111–120, Springer, San Bartolomé de Tirajana, Gran Canaria, Spain, November–December 2016.

[19] M. Nieto-Hidalgo, R. J. Valdivieso-Sarabia, J. Mora-Pascual, J. M. García-Chamizo, and F. J. Ferrández-Pastor, "Gait-A Database: dataset of subjects walking," 2017, http://hdl.handle.net/10045/70567.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 1310345, 11 pages
https://doi.org/10.1155/2017/1310345

Research Article

Gait Speed Measurement for Elderly Patients with Risk of Frailty

Xavier Ferre,[1,2] Elena Villalba-Mora,[1] Maria-Angeles Caballero-Mora,[3] Alberto Sanchez,[1] Williams Aguilera,[1] Nuria Garcia-Grossocordon,[3] Laura Nuñez-Jimenez,[3] Leocadio Rodríguez-Mañas,[3] Qin Liu,[2] and Francisco del Pozo-Guerrero[1,4]

[1]*Center for Biomedical Technology, Universidad Politécnica de Madrid, Madrid, Spain*
[2]*School of Software Engineering, Tongji University, Shanghai, China*
[3]*Servicio de Geriatría, Hospital Universitario de Getafe, Getafe, Madrid, Spain*
[4]*Biomedical Research Networking Center in Bioengineering Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain*

Correspondence should be addressed to Xavier Ferre; xavier.ferre@ctb.upm.es

Received 30 June 2017; Revised 22 October 2017; Accepted 29 October 2017; Published 31 December 2017

Academic Editor: Pino Caballero-Gil

The ageing of the population poses a threat to both public and private health and social systems. In the last 50 years, life expectancy has increased by an average of 20 years, and by the year 2050, life expectancy will exceed 90 years of age. However, quality of life in the last years of life is not guaranteed due to conditions such as functional decline and frailty, ultimately progressing to disability. Thus, the detection of such a condition in time is of utmost importance. This paper presents an ultrasonic sensor-based gait speed measurement device controlled via a mobile interface, which permits patients to self-assess physical performance. The system was developed and validated in an iterative process involving a total of 28 subjects (21 in the first round and 7 in the second one). After first evaluation at Hospital Universitario de Getafe, some technical problems arose whereas usability testing was well evaluated. The second version addressing the identified issues was technically validated at university premises with good and promising results. Future work envisages deployment of the system developed at subjects' homes to be remotely and unobtrusively monitored.

1. Introduction

According to the WHO (World Health Organization), in the last 50 years, life expectancy has increased by an average of 20 years, and if it continues increasing, by the year 2050, life expectancy will surpass 90 years of age [1]. Thus, elderly associated conditions and illnesses are more common today than they were in the past. For these reasons, major efforts are being undertaken investigating and analyzing the effects of aging, in conditions such as frailty, their impact on the cognitive and physical functions, and how to prevent these to improve the quality and duration of life. Frailty is defined as a state of increased vulnerability to adverse outcomes due to a reduction in the ability to respond to stressors [2]. The trajectory of frailty is usually characterized by a progression of functional decline that culminates in dependency and disability [2, 3].

The early detection of impeding disease is complex, and a clear algorithm and clinical-friendly screening tools for detection of frailty and/or disability are lacking. Living longer does not necessarily equate to healthy living. Development of early detection tools will permit intervention to prevent or delay the onset of frailty, thus preventing further disability.

Small changes in various health-related metrics could indicate the existence of an underlying condition in the patient. Such a condition might not present with visible symptoms until much later, when it is too late to take preventive measures and very costly to intervene, resulting in lengthy hospital stays and a higher risk for permanent institutionalization. Metrics such as the gait speed are used to measure a person's frailty level [4–6].

Gait speed deterioration is one of the most important indicators of decay in functional ability, with the risk of developing frailty, and later disability [7]. Early detection of any functional decay is critical to improve the quality of life of elderly people and reduce their level of dependency.

Gait speed is typically measured by health professionals in hospital or daycare settings, by manually measuring the time

the patient takes to walk between two stripes situated in the floor at different distances, depending on the tool used. In the case of the classical SPPB (short physical performance battery) test [8], the distance is 4 meters, but for Fried's criteria, this distance is 4.3 meters [2]. Healthcare professionals use a manual chronometer, which has a risk of inaccuracy and discrepancies between different professionals and different sets of measurements (inter- and intraobserver variability). Additionally, patients may spend several months without getting their gait speed measured, possibly worsening their functional condition to an irreversible state before any health professional can intervene.

This paper presents the design, construction, and validation of an ultrasonic sensor-based gait speed measurement device controlled via a mobile interface, which aims to empower patients in the unsupervised self-assessment of physical performance in both home and hospital settings and to make the measurement more reliable for health professionals.

Next section introduces related work in the subject. Section 3 presents the first version of the proposed device and its clinical evaluation in a hospital environment. Section 4 presents the modified version of the device and the technical validation of this second version. Finally, section 5 presents the conclusions gathered.

2. Related Work

Research on systems for gathering health-related metrics, and for measuring gait speed, can be divided into wearable- and non-wearable-based devices and sensor networks.

2.1. Wearable Devices. Most of the works related to monitoring and sensorization of medical analysis for elderly patients are based on the usage of wearable devices which gather the data and sometimes even perform simple measures. For example, Tirosh et al. proposed a sensorized sock equipped with pressure sensors to measure the variables for the gait test [9]. Adelsberger et al. proposed equipping a patient with inertial sensors to obtain measures for the gait analysis as well and send these data wirelessly via Bluetooth to a laptop computer for further analysis [10]. In the same line, there are many proposals that make use of different wearable sensors for data collection, such as the one from Morris and Paradiso [11] that advocates integrating the sensors to the patient's insoles and the one from Atallah et al. [12] that proposes to measure variables through a series of accelerometers placed on the patient's ear, among others.

Alternatively, Miura et al. proposed to use a mobile phone hung around the neck and fitted inside the user's jacket to measure walking characteristics, such as walking distance, time taken, speed, balance functions, body sways, and the number of steps taken. However, its accuracy for gait speed is too low for the SPPB test [13].

Finally, ultrasound sensors have been used to estimate the gait speed of subjects embedded with wearable devices; Weir and Childress proposed a device that combines ultrasound pulses and infrared technology to determine the real-time position of a user with respect to a base unit. Gait speed profile analysis provides parameters such as the cadence, step length, or step time variation [14].

The main shortcoming of the usage of these wearable devices that need to be attached to the patient in some way, is their intrusiveness. They may alter the conditions in which the analysis is performed, affecting users' behavior, and in some cases, they could be incompatible with some users' conditions. For example, some patients use orthopedic insoles on a daily basis, which makes it impossible to introduce a second insole to measure walking variables. The same issue can occur with socks or shoe-coupled devices as well as other nonwearable devices.

Some proposals use diverse ambient sensors for gathering vital variables. One of the most-used devices for this purpose is the Kinect®, developed by Microsoft, which groups a series of cameras and infrared sensors, allowing it to detect depth, record videos, and analyze movements and posture, among other features. Several research papers are based on this device, such as [15, 16], taking advantage of its resources to record the measures and then to model a series of significant parameters, such as longitude of the stride, walking speed, and walking direction, among other measures. These proposals are designed to work in the patient's home and collect all possible data for further analysis, obtaining extrainformation through modeling that could be important for the health professional. Kinect-based systems may raise privacy concerns, since it is composed of a camera recording users.

Other studies, like [17, 18], propose sensorized mats, based on a pressure sensor matrix, for obtaining these variables. These solutions are mostly restricted to use inside a health facility, operated by health professionals in a controlled environment. They are too costly and too difficult to calibrate for use in a home setting.

When considering in-house monitoring systems, there is an added difficulty in the analysis and evaluation of data because the measured events occur without monitoring, as opposed to the controlled environment of a medical office. When the patient is at home, there is no easy way of detecting when a walk begins and when it ends, if there are external factors affecting the speed and direction of the walk (e.g., if the patient is carrying something) or if there is some other variable affecting the measurement. Stone and Skubic minimize the "noise" introduced by these factors through the application of decision models, algorithms, and the gait variables to identify "valid" walks [16]. Van Den Broeck et al. proposed in their investigation a formula, using global coherence field (GCF), to estimate and remove the possible effects of the aforementioned noise [19].

2.2. Sensor Networks. Some research efforts have been focusing on behavioral monitoring, such as how to use sensor networks to help in the early detection of medical problems and to monitor adherence to doctors' recommendations regarding lifestyle and medication intake for chronic illness [20].

The University of Rochester built a smart medical home in their Center for Future Health, filled with different types of sensors, functioning as a laboratory, to test and develop various gadgets, sensors, and applications, that could be used as part of a health-integrated system [21]. Similarly, the Aware

Home in Georgia Tech, is a 3-storey, 5040 square feet facility designed to facilitate research in three main areas: health, entertainment, and sustainability, investigating how new technologies can impact the lives of people at home [22].

Implementing home sensors could give the elderly population the perfect balance between independence and proper monitoring. For example, Zouba et al. proposed the implementation of multiple sensors attached to furniture at home, combined with cameras for video analysis to monitor the behavior of elderly people performing everyday activities, such as cooking meals and eating [23]. Wood et al. took behavioral monitoring a step further by combining it with assisted living. They took data gathered from multiple heterogeneous sensors and fed them into a backend layer that analyzes and combines the data in order to produce decisions that could influence other devices, such as a power manager [24]. This system was thought to be extensible with the ability to add more sensors and user interfaces, which is a similar objective to the one proposed in this work. As with many other proposed solutions, it uses a proprietary query software language called SensQ, developed solely for this system which makes it harder for other consumers to adopt it. It also has a closed architecture which makes it complicated to integrate 3rd party external developments. These two limitations usually result in increased deployment costs and the risk of obsolescence.

3. UltraGaitSpeed System v1.0

For the measurement of the gait speed in a home setting, we have designed a lightweight and low-cost ultrasonic sensor-based system called UltraGaitSpeed. The sensor device connects via Bluetooth with an Android mobile phone that logs the measured data and allows the user to control the system. When an Internet connection is available, the data are sent to a Cloud server.

Our approach to gait speed measurement aims to offer a lightweight and unobtrusive solution that liberates the user from wearing any kind of gadgets and takes advantage of a very powerful device (modern smartphones) that has a tremendous processing power and good quality/price ratio.

A strip of light sensors with an Arduino board can be easily deployed in a house and can provide the necessary information to monitor decay in gait speed, possibly indicating more serious health problems that require immediate assistance. The gathered information is uploaded to a server and securely connected to the HIS (hospital information system), so physicians can monitor patient evolution, and alarms can be raised in a clinical setting when the condition of a patient rapidly deteriorates.

3.1. System Design. We chose HC-SR04 ultrasonic sensors for their lightweight and low-cost characteristics. This sensor uses sonar, sending a signal from one side and receiving it on the other, to measure the distance of an object in front of it by subtracting the time between both signals. This means that we can detect when an object moves in front of the sensor by tracking the distance variations between measures. Since these sensors emit an ultrasonic sound wave and sound waves dissipate in a cone-shaped way, we need to perform some tests with the sensors to make a decision about the number of sensors and to establish a minimum distance between sensors to avoid interference (Figure 1).

After extensive testing, we determined that 35 centimeters is the minimum distance between sensors for them to work without interference and that 5 sensors provided enough accuracy since additional sensors did not improve the measurements (Figure 2).

The 5 sensors are connected to an Arduino UNO board and set up with 1-meter separation between them along a flexible EVA foam strip. The resulting 4-meter strip can serve to reproduce the SPPB gait speed test used by health professionals.

In order to cover the sensors and protect the cables, cases were printed in 3D for every sensor and its connections (Figure 3). This provides the required robustness for a solution to be installed in a home setting.

The device runs a simple software that loops through the connected sensors, sends a signal, and listens for the response, which is stored in a local variable. After looping through all the sensors, it sends via Bluetooth the data gathered from the sensors in JSON format, with the number of milliseconds relative to the system start as a timestamp.

The mobile application allows the user to calibrate the sensors, start the measurement, and get data results on the screen. It also stores measurements in an internal log, which is uploaded to a server when an Internet connection is available.

This first version of the UltraGaitSpeed system works according to the flowchart shown in Figure 4. The application guides the user in all the steps required for system operation, namely, Bluetooth enabling, sensor calibration, step measurement, and final result display.

Bluetooth connection is enabled in two steps. First, the application checks the activation of Bluetooth in the device and then connects with the Arduino Board. Next, the sensor calibration starts, using a series of iterations to make sure that stable distances are measured by each sensor. These distances are the baseline for the later measurements. If there is movement in front of any sensor, the calibration will fail. If there is a failure in any of these steps, the application displays an error message to the user.

When calibration finishes, the system starts measuring gait speed. Measurement is achieved by means of detection events. When a user is detected by a sensor, the distance measured deviates from the sensor baseline measure and a timestamp is logged. Finally, when every sensor has detected the patient, the passing time at each sensor and the overall time are calculated by subtraction.

By pressing the Restart button on the screen, the process can start again for a new measurement.

3.2. Evaluation. Even if our system is meant to be used by patients and informal caregivers in an extrahospital setting, we need to validate that the system gathers gait speed measurements as reliably as the manual measurement by health professionals in medical facilities.

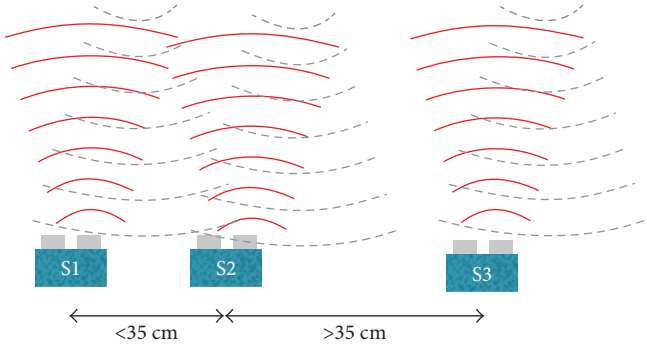


Figure 1: Minimum distance between sensors.

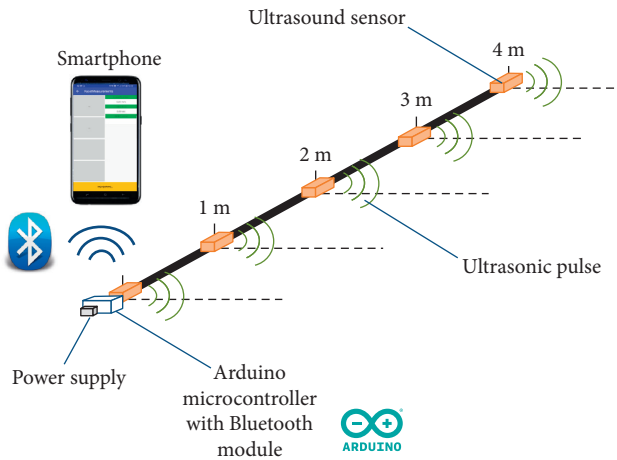


Figure 2: Equidistant sensor placement in the strip.

For that purpose, we carried out an experiment with older patients attending the Day Hospital in the Geriatric Service at the Hospital Universitario de Getafe (HUG). The objective is to validate the accuracy and feasibility of the proposed measurement solution as compared to the current manual practice.

3.2.1. Experiment Design. Our hypothesis is that the usage of the proposed ultrasound sensor-based solution allows the conduction, in a precise way, of the SPPB gait analysis, with a difference threshold of ±0.5 seconds compared to the manual measures, which is the current practice in Geriatric Service at HUG.

The inclusion criteria for the patients to take part in the experiment were as follows:

(i) Older than 70 years

(ii) Patient of the Geriatric Service at HUG.

The exclusion criteria were as follows:

(i) Cognitive impairment: minimental test result lower than 20

(ii) High degree of disability: Barthel index lower than 60.

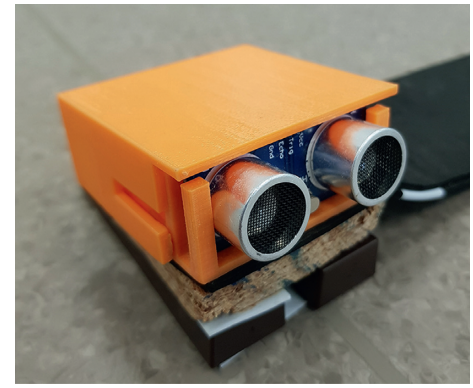


Figure 3: 3D printed case for covering sensors and connections.

The patients were randomly selected from those that come to the Day Hospital, with at least 2 patients selected in each one of the following groups:

(i) Patients that do not need any technical help for walking

(ii) Patients that need a cane for walking

(iii) Patients that need a walker for walking.

Two separate tests were conducted with each patient. First, the gait test as described in point 2 of the SPPB analysis was used to measure the speed of the walk, having the patient standing right on the start line of the 4-meter strip, including acceleration speed (static start). Second, an ordinary gait test was performed, having the patient start a meter before the starting line (dynamic start) in order to exclude acceleration from the measure.

The experiment was designed to simultaneously measure using both the current manual method and the proposed system. The manual method is done with the patient walking between two green lines while the health professionals measure with a stopwatch. For each of the tests, three measurements were taken in total: two separate manual measures and one with the proposed solution, all of them done by health professionals experienced in measuring gait speed. None of the health professionals involved in taking the measures could see the measures taken by the others.

Forms for the manual measure results were prepared as well as an SUS (System Usability Scale) [25] satisfaction questionnaire and an impressions questionnaire to gather information about the system usability from the health professional using our proposed system in the experiment.

The current protocol for gait speed measurement for the SPPB test was followed: measuring two walks for each patient and taking as valid the best of both measures.

The Ethical Committee of the University Hospital of Getafe approved the protocol.

3.2.2. Experiment Execution. On the day of the experiment, we set up the data gathering device and the mobile application to conduct the tests. The setting up of the gathering device was a very simple procedure, taking less than 5 minutes.

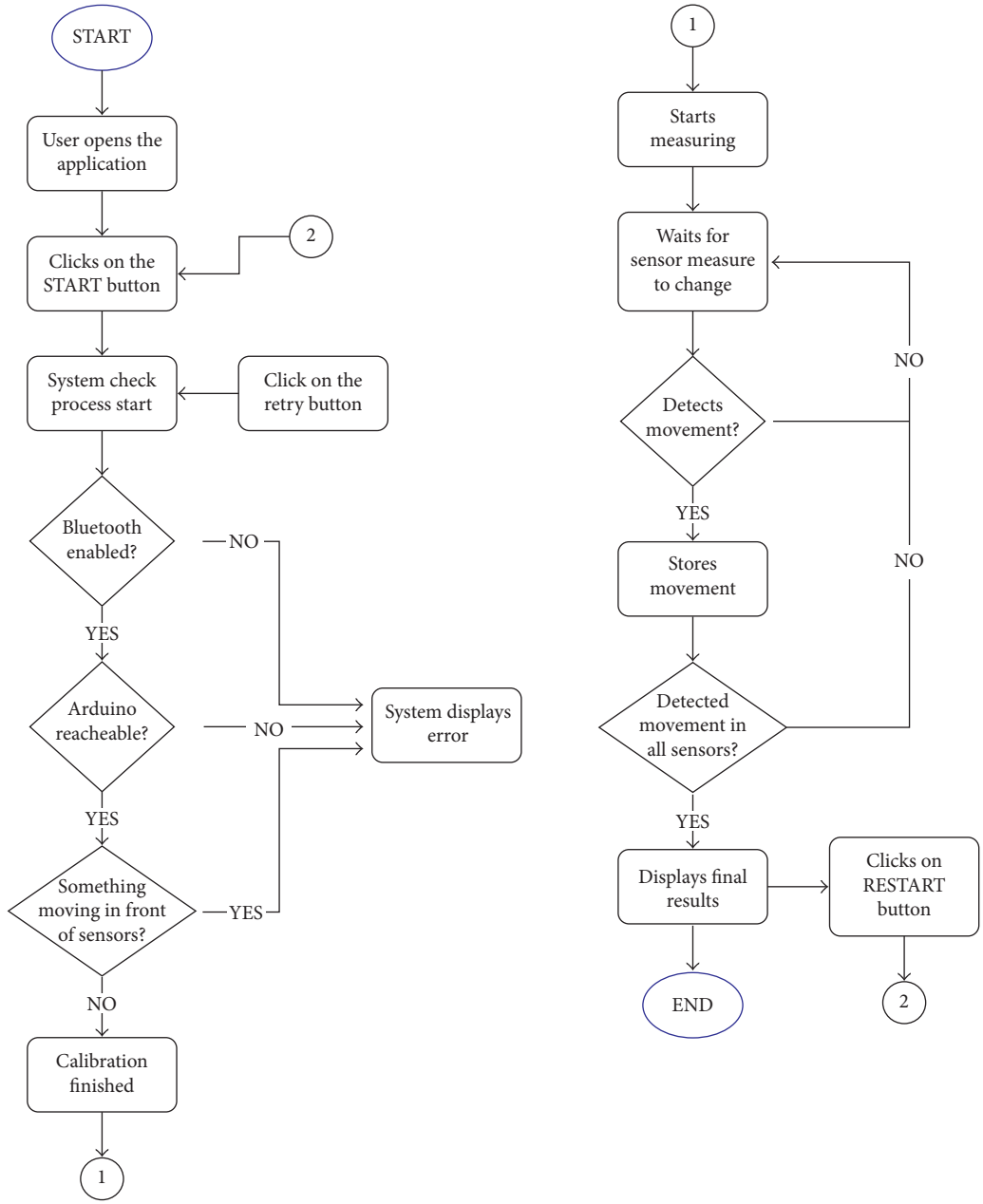


Figure 4: Algorithm for UltraGaitSpeed v1.0.

Three health professionals volunteered to participate in the experiment as data gatherers, among whom we randomly distributed the roles for manual collection and usage of the application. After selecting the professionals, we gave them a brief explanation of the experiment and their tasks and provided them with the individual data collection forms. A mobile device that had the application already installed was provided to the health professional for taking measurements.

The experiment was conducted as follows:

(1) After selecting those patients that met the inclusion criteria and that expressed willingness to participate (by signing the informed consent), the health professional proceeded to explain what the test consisted of.

(2) The subjects were asked to carry out the two tests (the static and dynamic start). Their gait speed was measured by health professionals. One professional measured gait speed automatically with the proposed system, and the two other professionals performed the test manually using the current method (stopwatch). Manual results were collected in an ad hoc paper form.

(3) After finishing with all the patients of the study, we gathered the paper forms from the health professionals and asked the one that was using the application to complete the satisfaction and impressions questionnaires.

After carrying out the experiment, an analysis of the data was conducted to ensure the feasibility of the system through the analysis of the following variables:

(i) Number of measures that ended correctly

(ii) Number of measures that had to be repeated

(iii) Time for the preparation of each test

(iv) Deviation related to the control measure.

The experiment was conducted with 21 voluntary patients that complied with the inclusion and exclusion criteria. Their average age was 81 years, with the youngest patient being 71 and the oldest 90. The number of patients with technical aids, such as canes and walkers, was 7 from the total of 21 patients that were included in the study.

The experiment with patients went smoothly with no need for further intervention from the researchers. There were slight problems with the calibration in patients who use walking canes, which were resolved after rebooting the system.

3.2.3. Results. After finishing the experiment, we proceeded with the extraction of data gathered from the forms and the questionnaires completed by the health professional using the application.

First, we merged the data gathered into one single spreadsheet, grouped by a type of measure (stopwatch 1 and 2 and application). After this, a series of data treatments were applied to the data as follows:

(1) We selected as final measure for each subject the lower measure between the two that were taken for each one of the tests (the dynamic and static start), as is usually done by health professionals when performing this test.

(2) With the final measure for each of the subjects, we calculated the average between the manual measures taken with the stopwatch to compare those measures with the ones obtained from the application.

(3) We then proceeded to calculate the difference module of the measures obtained with the manual procedure, and those from the application, and all the combinations between them:

   (a) Between manual measure 1 and manual measure 2
   (b) Between the average of the manual measures and the application
   (c) Between manual measure 1 and the application
   (d) Between manual measure 2 and the application.

(4) Having these measures, we proceeded to calculate statistical variables of the differences. These are relevant in order to measure the degree to which the application was able to get correct measures in the experiment. The variables obtained were the mean, the median, the mode, and the truncated mean.

(5) We decided to calculate the truncated mean of the differences with a threshold of 10%. This allowed us to remove 10% of outliers from the calculation of the mean (which accounted for 2 measures), thus omitting occasional errors that the application might have had.

After data processing, we obtained the results detailed in Table 1 for the case of the measure with no acceleration and in Table 2 for the case of the measure including the patient's acceleration. Overall relevant variables of the experiment are shown in Table 3.

Regarding the satisfaction questionnaire, the responses produced an SUS score of 72.5, which is slightly over the average of usability scores taken as a reference for an SUS score of 68.5. This shows that there is satisfaction in the user but not to an exceptional level.

The open impressions questionnaire shows a favorable perception from the health professional, and it offers several relevant suggestions for improvement:

(i) To optimize calibration and to not require the patient to wait, standing without movement while the system calibrates. Elderly patients have difficulty standing for some time, since they get easily tired. In some cases, there is a risk that they lose their balance

(ii) To avoid having to perform a calibration every time, a measure is started. The system could calibrate only the first time, providing a button to recalibrate when the health professional sees fit

(iii) To extract additional variables from the sensors, such as direction of the walk, variability of the speed, stride length, step length, and others. For the implementation of this suggestion, we will need to review and perform a deeper analysis of the gathered data, in order to determine which of these proposed new variables can be obtained from the current sensors.

3.2.4. Discussion. These results show the feasibility of deploying our proposed system in a clinical setting and having health professionals use it. But the reliability of the measures falls below the threshold established of a maximum difference of 0.50 seconds between the manual measure and the automatic measure. Considering that all the measures fall in the interval 2.15–9.47 seconds, average differences of 0.64 or 0.67 are too high.

Further detailed study of the results with health professionals at HUG unveiled that the algorithm used in UltraGaitSpeed v1.0 for measuring gait speed did not exactly comply with the measuring criteria used in classical gait speed tests, as routinely performed at HUG. Both for the static and dynamic start gait speed tests, the event triggering the start and finish of measuring is the moment when the last heel of the subject reaches the mark on the floor, while the algorithm implemented in v1.0 of our solution starts measuring when the sensor detects the first foot. This is why the presence of walking aids produced some errors.

Regarding calibration, the system failed several times for the static start gait speed tests. In this case, the subject stands right in front on the starting line, with the tip of his foot on it, and waits for the command to go. Even the smallest variation in the distance detected at the starting line, due to subjects'

Table 1: Experiment results for the dynamic start measure with UltraGaitSpeed v1.0.

	Differences (seconds)			
	Stopwatch 1 and stopwatch 2	Stopwatch avg. and UltraGaitSpeed v1.0	Stopwatch 1 and UltraGaitSpeed v1.0	Stopwatch 2 and UltraGaitSpeed v1.0
Mean	0.13	0.66	0.66	0.67
Median	0.06	0.71	0.70	0.63
Mode	0.03	No mode present	0.75	1.13
Truncated mean (10%)	0.10	0.62	0.62	0.62
Average relative error with respect to manual watch average time (%)	2.8%	13.6%	13.7%	14.0%
Standard deviation	0.13	0.45	0.42	0.50

Table 2: Experiment results for the static start measure with UltraGaitSpeed v1.0.

	Differences (seconds)			
	Stopwatch 1 and stopwatch 2	Stopwatch avg. and UltraGaitSpeed v1.0	Stopwatch 1 and UltraGaitSpeed v1.0	Stop watch 2 and UltraGaitSpeed v1.0
Mean	0.20	0.64	0.64	0.64
Median	0.13	0.45	0.38	0.49
Mode	0.28	No mode present	0.13	0.47
Truncated mean (10%)	0.16	0.55	0.53	0.56
Average relative error with respect to manual watch average time (%)	3.7%	11.9%	11.9%	11.9%
Standard deviation	0.20	0.57	0.63	0.53

or walking aids' movement, triggered the timing and produced irregular measurements.

Additionally, the system requires that the test is carried out only in one direction, while software could detect the walking direction and thus capture two measurements per patient without requiring him or her to walk back to the starting point. The patient could just turn around after the first measurement and walk in the other direction for the second measure.

Table 3: Variables obtained from the experiment with UltraGaitSpeed v1.0.

Variable	Result
Time for the preparation of each test	Mean of 4 minutes
Deviation related to the control measure	Mean of 0.58 seconds
Maximum time measured	9.47 seconds
Minimum time measured	2.15 seconds

4. UltraGaitSpeed v2.0

In order to overcome the main problems identified for v1.0, we designed a new algorithm (Figure 5) and a new version for the mobile app (Figure 6), the UltraGaitSpeed v2.0.

The new algorithm considers the last detection event in a sensor to be the moment to consider for user passing such a sensor (Figure 7). This better fits the moment when the patient's heel leaves the sensor detection field. This moment is used for both establishing the start and finish of the measure, mirroring the way the health professionals define the start and finish time when performing this test. Likewise, this criterion determines partial timing.

The distance between sensors is fixed and known, and the detection field angle is the same for all of them. Therefore, if we assume that the patient will walk straight, the point where the last foot abandons each detection field will be separated by the same distance as the sensors themselves.

So, with v2.0 of our device, we can measure the time spent moving from sensor A to sensor B.

Figure 5 depicts the new algorithm designed for v2.0 of our system. When one of the sensors detects the user, it is activated and waits for the field-of-view exit event, at which point the timestamp is registered. In order to make sure that the exit event was actually the last part of the body to pass (and not, e.g., just the first foot or a walking cane), the measurement for each segment of the gait speed test is not confirmed until the exit event is produced in the next sensor in the strip. (The last sensor has a timeout for confirmation, since it has no other sensor next to it.) In the case of a new detection event before confirmation in the following sensor, the sensor waits for the expected exit event and then the timestamp is reset.

This new algorithm entails the following advantages:

(1) It replicates the measuring criteria used in the classical gait speed tests carried out by professionals in clinical settings.

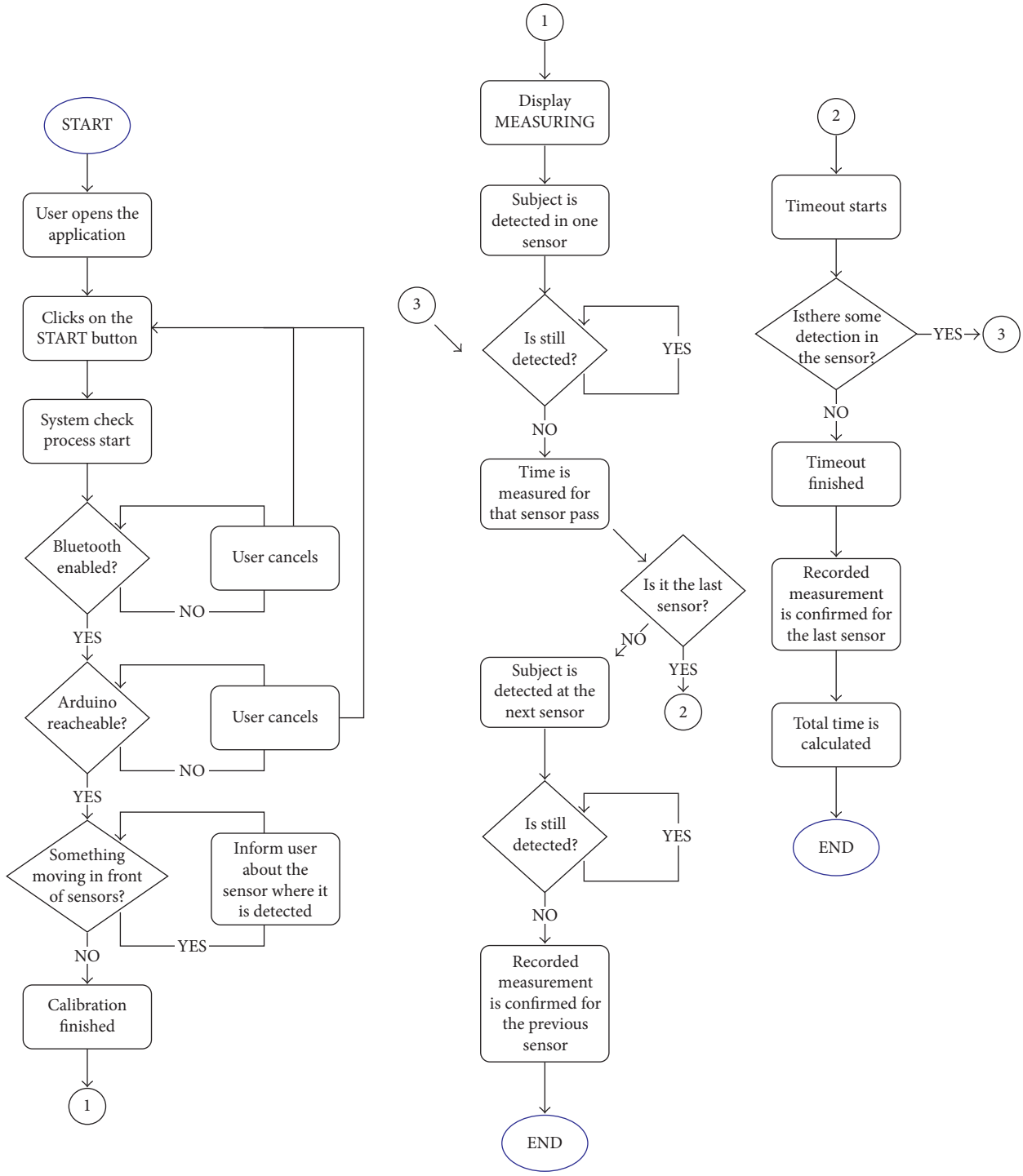


Figure 5: Algorithm for UltraGaitSpeed v2.0.

(2) Start time is automatically triggered when the patient leaves the first sensor, instead of requiring an action in the mobile app.

(3) Calibration does not require that the patient waits motionless just before the first sensor. It is carried out with the patient positioned behind or far away from the sensor strip, and the measuring will automatically start when the patient goes through the first sensor. We will thus avoid errors arising from the small distance variations while patients wait for the calibration to end.

4.1. Evaluation. We have evaluated this new version of our proposal with seven users between 68 and 79 years of age, six of them with mobility problems and two requiring a cane.

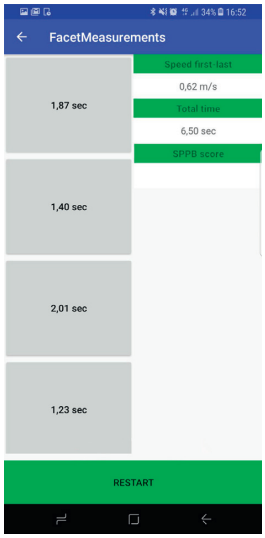


Figure 6: Screen capture of the mobile app for UltraGaitSpeed v2.0.

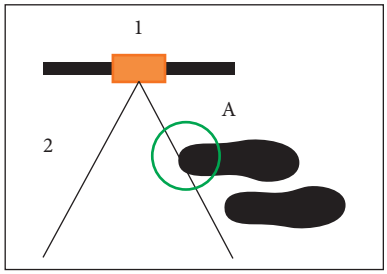


Figure 7: Detection moment for the user passing through a sensor.

We have deployed UltraGaitSpeed v2.0 at CTB with indications of the place for subjects to position themselves for both the dynamic and the static start tests. We have briefly explained how the system works, and how it is controlled from the mobile app to the subject, in the same way as we foresee system installers doing when the system is actually used in a home setting.

Each participant is asked to measure two walks with a static start and two more with a dynamic start, in total providing four measures to the system.

A biomedical engineer trained at HUG on the SPPB test measures manually with a stopwatch the time spent by the user between the two extremes of the sensor strip (4 meters). The measurer followed the criteria used in regular clinical practice. Timing must be recorded when the last heel of the patient passes over the line that marks the start and finish of the four-meter walk.

4.1.1. Results. The results obtained for timing were analyzed replicating the protocol of UltraGaitSpeed v1.0 evaluation but having only one stopwatch measurement as a reference for comparison.

The results obtained are presented in Tables 4 and 5.

Table 4: Experiment results for the static start measure with UltraGaitSpeed v2.0.

Differences (seconds) UltraGaitSpeed v2.0 and stopwatch	
Mean	0.35
Median	0.29
Mode	None
Truncated mean (10%)	0.35
Average relative error with respect to manual watch mean time (%)	7.39%
Standard deviation	0.22

Table 5: Experiment results for the dynamic start measure with UltraGaitSpeed v2.0.

Differences (seconds) UltraGaitSpeed v2.0 and stopwatch	
Mean	0.15
Median	0.19
Mode	None
Truncated mean (10%)	0.15
Average relative error with respect to manual watch mean time (%)	4.13%
Standard deviation	0.10

The new algorithm demonstrated a greater reliability for measurement than the one implemented in the previous version of the device. Both the absolute and relative errors between the stopwatch and UltraGaitSpeed v2.0 have been reduced. In fact, the average absolute error is smaller than the established accepted error of 0.5 seconds. Moreover, the standard deviation of the error is smaller compared with the first version, which suggests the error of this version falls in a smaller range, and therefore is more predictable.

The timing problem associated with the use of walking aids has been solved in this new version. The trials with two patients who use walking aids demonstrated the reliability of UltraGaitSpeed v2.0. The average relative error for their tests was smaller than the overall (3.95%), and the measurer did not report any incident related to this issue.

Observation of subjects while using the system and open conversation with them have provided insight about some usability problems that need to be addressed as follows:

(i) The word "measuring..." conveys the idea that the time watch is already ticking, while it actually means that the sensors have been calibrated and are already working to detect whatever passes by.

(ii) Error and warning messages, such as the one appearing when the timeout for the last sensor goes off, need to be reconsidered to speak the user's language, and some of them should not even be shown unless the app is in developer mode.

Despite the room for improvement, UltraGaitSpeed v2.0 was accepted by the users. They completed the

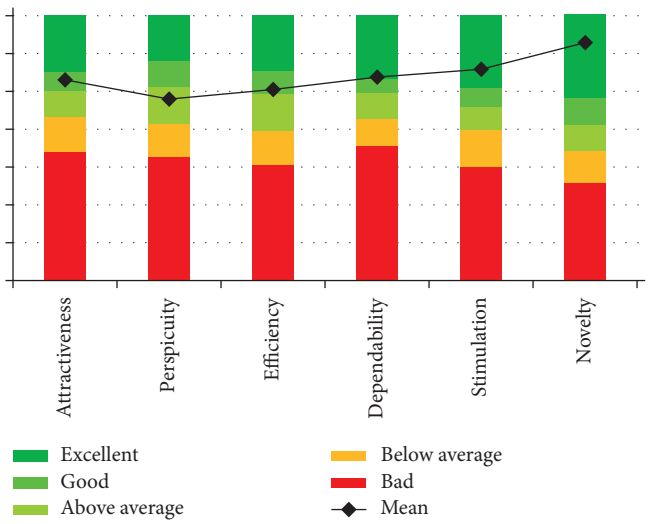


Figure 8: Results for UEQ.

System Usability Scale (SUS) and User Experience Questionnaire (UEQ) [26] after performing the gait speed tests, giving an average SUS score for UltraGaitSpeed v2.0 of 73.5.

Studies have demonstrated that the average SUS score is 68. Therefore, we have shown that our system is more usable than the benchmark and is in the second quartile [27].

In the case of the UEQ, we obtained excellent results in Novelty and Stimulation, good results in Attractiveness, Efficiency, and Dependability, and results above the average in Perspicuity (Figure 8).

4.1.2. Discussion. Despite the small size of the study population, results suggest that UltraGaitSpeed v2.0 may be used in unsupervised environments for evaluating the gait speed of patients, from a technical point of view.

Additionally, the new algorithm prevents one of the more recurrent errors of UltraGaitSpeed v1.0. The initial algorithm triggered incorrect start of timing when small movements of patients waiting to begin the static start test were detected as actual movement.

With these results, we confirm that the new design eliminates the error in the start of each measure, by having the measuring algorithm consider the moment when the user leaves each field of detection as the time of passing.

The validation of UltraGaitSpeed v2.0 has shown the system feasibility in its usage by the intended audience: elderly people. Participants in the validation have shown a favorable opinion of the usefulness and usability of the system, even though they suggested necessary improvements.

Further validation is necessary in the intended final environment: patients' homes. A longitudinal study is necessary to better understand how the system can be regularly used by its intended users.

5. Conclusions

We have presented two versions of UltraGaitSpeed (v1.0 and v2.0), a sensor-based system controlled through a mobile app for measuring gait speed, aimed to be used by patients themselves or their caregivers.

Through the two-cycle validation and redesign process, the proposed system has shown its usefulness to automatically measure gait speed through a low-cost and lightweight solution that can be easily deployed in any medical facility, in a daycare center, or even in a private home. As an Android app controls the system, this allows its usage in the patient's or informal caregiver's own mobile phone, thus further reducing the cost of the integrated solution.

Even if the validation has been carried out in a controlled environment with health professionals or researchers, the proposed system can be part of an ambient-assisted living network of sensors, that contributes to a better quality of life and earlier detection of deteriorating conditions that require professional medical help.

The usability level of the system needs to be improved with a new redesign-evaluation loop, where acceptability and usability issues are resolved.

The next step in our research is to validate the system in a hospital setting like the one used in the validation of the first version and a longitudinal study carried out in in-home settings. In these studies, we will inquire about the most appropriate amount of sensors and length of the strip. The aim is to reduce the strip to a length of 2 meters, facilitating its reliable usage in home environments with limited space.

The system will be extended to address other SPPB test parts and other elements from the comprehensive geriatric assessment like the chair stand test (measuring how long the patient takes to sit down in a chair and stand up again five times), involuntary loss of weight, and the items of the Linda Fried criteria of frailty. We will provide these elements either through sensor-based systems like the one presented in this paper or through an extension of the mobile app.

At a subsequent stage, the aim is to create a system based on UltraGaitSpeed that measures gait speed unobtrusively, without the need for the user to actively start the measuring. If the sensor strip can be inconspicuously installed in a corridor wall and measure gait speed every time the user passes by, it will provide much richer information without the added stress of following instructions and manipulating a mobile phone.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the EIT Health in the FACET (FrAilty Care and wEll funcTion) project and by funds from CIBERfes (FEDER). The authors would like to thank all the

volunteers that helped them to validate their solution acting as test subjects and to thank the reviewers for their insightful comments.

References

[1] D. M. Ediev, "Life expectancy in developed countries is higher than conventionally estimated. Implications from improved measurement of human longevity," *Journal of Population Ageing*, vol. 4, no. 1–2, pp. 5–32, 2011.

[2] L. P. Fried, C. M. Tangen, J. Walston et al., "Frailty in older adults: evidence for a phenotype," *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 56, no. 3, pp. M146–M157, 2001.

[3] T. M. Gill, E. A. Gahbauer, L. Han, and H. G. Allore, "Trajectories of disability in the last year of life," *New England Journal of Medicine*, vol. 362, no. 13, pp. 1173–1180, 2010.

[4] M. Runge and G. Hunter, "Determinants of musculoskeletal frailty and the risk of falls in old age," *Journal of Musculoskeletal and Neuronal Interactions*, vol. 6, no. 2, pp. 167–173, 2006.

[5] J. M. Hausdorff, D. A. Rios, and H. K. Edelberg, "Gait variability and fall risk in community-living older adults: a 1-year prospective study," *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 8, pp. 1050–1056, 2001.

[6] J. M. Hausdorff, A. Schweiger, T. Herman, G. Yogev-Seligmann, and N. Giladi, "Dual-task decrements in gait: contributing factors among healthy older adults," *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 63, no. 12, pp. 1335–1343, 2008.

[7] E. Villalba Mora, R. Petidier-Torregrossa, C. Alonso-Bouzon, J. A. Carnicero-Carreño, and L. Rodríguez Mañas, "Early detection of Heart Failure exacerbation by telemonitoring in old people," *International Journal of Integrated Care*, vol. 15, no. 5, pp. 116972–116974, 2015.

[8] J. M. Guralnik, E. M. Simonsick, L. Ferrucci et al., "A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission," *Journal of Gerontology*, vol. 49, no. 2, pp. M85–M94, 1994.

[9] O. Tirosh, R. Begg, E. Passmore, and N. Knopp-Steinberg, "Wearable textile sensor sock for gait analysis," in *Proceedings of the 2013 Seventh International Conference on Sensing Technology (ICST)*, pp. 618–622, IEEE, Wellington, New Zealand, December 2013.

[10] R. Adelsberger, N. Theill, V. Schumacher, B. Arnrich, and G. Tröster, "One IMU is sufficient: a study evaluating effects of dual-tasks on gait in elderly people," in *Proceedings of the International Conference on Wireless Mobile Communication and Healthcare*, pp. 51–60, Springer, Paris, France, November 2012.

[11] S. J. Morris and J. A. Paradiso, "Shoe-integrated sensor system for wireless gait analysis and real-time feedback," in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference 2002 (Engineering in Medicine and Biology, 2002)*, vol. 3, pp. 2468-2469, IEEE, Houston, TX, USA, October 2002.

[12] L. Atallah, O. Aziz, B. Lo, and G. Z. Yang, "Detecting walking gait impairment with an ear-worn sensor," in *Sixth International Workshop on Wearable and Implantable Body Sensor Networks, 2009 (BSN 2009)*, pp. 175–180, IEEE, Berkeley, CA, USA, June 2009.

[13] T. Miura, K. I. Yabu, A. Hiyama, N. Inamura, M. Hirose, and T. Ifukube, "Smartphone-based gait measurement application for exercise and its effects on the lifestyle of senior citizens," in *Proceedings of the Human-Computer Interaction (INTERACT 2015)*, pp. 80–98, Springer, Bamberg, Germany, September 2015.

[14] R. F. Weir and D. S. Childress, "A portable, real–time, clinical gait velocity analysis system," *IEEE Transactions on Rehabilitation Engineering*, vol. 5, no. 4, pp. 310–321, 1997.

[15] M. Rantz, M. Skubic, C. Abbott, Y. Pak, E. E. Stone, and S. J. Miller, "Automated fall risk assessment and detection in the home: a preliminary investigation," in *Proceedings of the AAAI Fall Symposium: Artificial Intelligence for Gerontechnology*, Arlington, VA, USA, November 2012.

[16] E. E. Stone and M. Skubic, "Mapping Kinect-based in-home gait speed to TUG time: a methodology to facilitate clinical interpretation," in *Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pp. 57–64, IEEE, Venice, Italy, May 2013.

[17] A. L. McDonough, M. Batavia, F. C. Chen, S. Kwon, and J. Ziai, "The validity and reliability of the GAITRite system's measurements: a preliminary evaluation," *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 3, pp. 419–425, 2001.

[18] D. Gouwanda, N. Senanayake, M. Marasinghe et al., "Real time force sensing mat for human gait analysis," *International Journal of Mechanical and Mechatronics Engineering*, vol. 2, no. 3, pp. 349–354, 2008.

[19] B. Van Den Broeck, L. Vuegen, H. Van Hamme, M. Moonen, P. Karsmakers, and B. Vanrumste, "Footstep localization based on in-home microphone-array signals," in *Proceedings of the 12th Biennial European Conference of the Association for the Advancement of Assistive Technology in Europe (AAATE 2013)*, vol. 33, p. 90, IOS Press, Amsterdam, Netherlands, September 2013.

[20] A. Helal, D. J. Cook, and M. Schmalz, "Smart home-based health platform for behavioral monitoring and alteration of diabetes patients," *Journal of Diabetes Science and Technology*, vol. 3, no. 1, pp. 141–148, 2009.

[21] J. Marsh, "House calls," in *Rochester Review*, vol. 64, no. 3, pp. 22–26, University of Rochester, Rochester, NY, USA, 2002, https://www.rochester.edu/pr/Review/V64N3/feature2.html.

[22] Georgia Tech, *Aware Home Research Initiative*, Georgia Institute of Technology, Atlanta, Georgia, 2017, http://www.awarehome.gatech.edu.

[23] N. Zouba, F. Bremond, and M. Thonnat, "An activity monitoring system for real elderly at home: validation study," in *Proceedings of the 2010 Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 278–285, IEEE, Boston, MA, USA, September 2010.

[24] A. Wood, J. Stankovic, G. Virone et al., "Context-aware wireless sensor networks for assisted living and residential monitoring," *IEEE Network*, vol. 22, no. 4, pp. 26–33, 2008.

[25] J. Brooke, "SUS-A quick and dirty usability scale," *Usability Evaluation in Industry*, vol. 189, no. 194, pp. 4–7, 1996.

[26] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in *Symposium of the Austrian HCI and Usability Engineering Group*, pp. 63–76, Springer, Graz, Austria, November 2008.

[27] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: adding an adjective rating scale," *Journal of Usability Studies*, vol. 4, no. 3, pp. 114–123, 2009.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 5253613, 10 pages
https://doi.org/10.1155/2017/5253613

Research Article

Massage Therapy of the Back Using a Real-Time Haptic-Enhanced Telerehabilitation System

Cristina Ramírez-Fernández,[1] Victoria Meza-Kubo,[2] Eloísa García-Canseco,[2] Alberto L. Morán,[2] Oliver Pabloff,[2] David Bonilla,[2] and Nirvana Green[2]

[1]*Instituto Tecnológico de Ensenada, Ensenada, BC, Mexico*
[2]*Facultad de Ciencias, Universidad Autónoma de Baja California, Ensenada, BC, Mexico*

Correspondence should be addressed to Cristina Ramírez-Fernández; cramirezf@gmail.com

Received 29 July 2017; Accepted 15 November 2017; Published 31 December 2017

Academic Editor: Pino Caballero-Gil

We present the usability evaluation of a haptic-enhanced telerehabilitation system for massage therapy of the back using the Vybe haptic gaming pad and the gesture sensor LEAP motion controller. The evaluated system includes features that allow for (i) administering online therapy programs, (ii) providing self-adjustable and safety treatment of back massages using a virtual environment, and (iii) saving and replaying massage sessions according to a patient's therapy program. The usability evaluation with 25 older adults and 10 specialists suggests that the haptic telerehabilitation system is perceived with high usability and pleasurable user experience, while providing personalized intensity of haptic therapy in a supervised, real-time, and secure way to treat the patient. Moreover, the specialists totally agree that the system design features, such as save and play, and delimiting therapy zones are the most important for back massage therapy, while the features of regulating feedback intensity and providing/receiving a massage remotely are also important. Finally, based on their comments, five design insights aiming at improving the current version of the system were generated.

1. Introduction

Due to a combination of demographic changes and the lack of resources in the field of public health and technology improvement, the development of new rehabilitative practices seems mandatory to build sustainable models for rehabilitation from the clinical, organizational, and economic perspectives [1]. In recent years, haptic feedback has proven to enhance user experience in telerehabilitation [2]. However, current telerehabilitation systems have yet to exploit the richness of haptic modality within its content. The lack of haptic methods that provide real-time rehabilitation in a supervised and remote way inhibits the attention of patients that require massage treatment of the back. Massage therapy has become one of the most popular complementary and alternative medical (CAM) therapies for back pain, the condition for which CAM therapies are most commonly used [3]. A massage is defined as "a mechanical manipulation of body tissues with rhythmical pressure and stroking for the purpose of promoting health and wellbeing" [4]. Despite the availability of novel mechanical support tools for back massage intervention, for example, [5–7], massage therapy is still provided in a traditional way, where both the patient and the therapist need to be in the same place [8].

In [9], we introduced a novel real-time haptic-enhanced telerehabilitation system for massage therapy that consists of a web application and a local virtual environment (VE) in which the interaction is performed using the Leap Motion Controller (LMC) gesture sensor [10] and the Vybe haptic gaming pad [11]. The system, called GoodVybesConnect, allows the therapy input parameters to be individualized and calibrated according to the patient's characteristics. It also allows the execution of the therapy to be dependent on the therapist's hand movement in the VE, so that multimodal feedback (i.e., visual, audible, and vibrotactile feedback) is sent to patients in real time depending on the availability of the

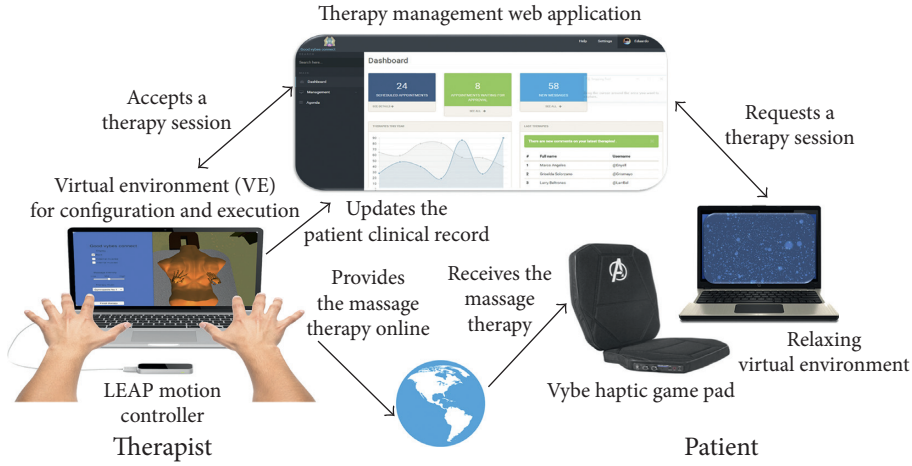


Figure 1: The GoodVybesConnect architecture.

therapist. Otherwise, the patient might play at home a massage session previously scheduled by the specialist. Finally, therapy results are automatically updated in the patient's clinical record.

To assess the GoodVybesConnect current prototype and to identify some design insights that might help improving the current version of the system, we carried out two evaluations: the objective of the first evaluation was to identify the user experience of potential patients during a massage session, while the second one aimed at knowing the specialists' (i.e., therapists and psychologists) opinions regarding their usability perception of the application and their user experience while participating as patients and therapists.

2. Related Work

According to the literature, there are various forms of intervention for back treatment, including physical therapy, acupuncture, chiropractic treatment, and massage therapy, among others [12]. Massage therapy has been shown to affect both structure and function of the musculoskeletal system by promoting a relaxation response, decreasing muscle tension, and decreasing tonic muscle contractions [13–16]. In recent years, the integration of mechanical instruments to support the implementation of massage therapies such as massage chairs [5–7], automatic massage systems [6], or robotic systems [17–21] has benefited therapists and their patients.

The massage chair allows different massage techniques, for example, pressure point massage, roll stretch massage, and beat massage [5]. In some cases, depending on the patient's physiological variables being measured (e.g., mental condition, blood pressure, heart rate, and stress, among others), the specialist selects the massage technique that she/he considers the most appropriate [22–24]. There also exist automatic massage systems such as medical devices that combine basic principles of mechanical massage, thermotherapy, acupressure, infrared therapy, and moxibustion. For instance, the device described in [6] is capable of releasing modulated and controlled thermomechanical energy on the patient following a programmable and fully reproducible automatic treatment program selected by the operator. Another example of automatic massage systems is robotic systems that use massage patterns, for example, robots that provide maxillofacial massages [17], massage of the back [20, 21, 25], or feet massage [18]. Nevertheless, the implementation of massage therapy with these mechanical instruments needs "hard-wired" programmed routines, local supervision of the therapist, and manual therapy management programs [26]. Moreover, these procedures also require the patient to be in the same place as the therapist, and they are mainly carried out in the traditional way using nonmechanical instruments [3, 12].

In comparison to the aforementioned references, our real-time haptic-enhanced telerehabilitation system [9] is innovative given that it (i) proposes a method to provide real-time back rehabilitation in a supervised and remote way, (ii) provides a web application for managing therapy programs, (iii) combines emerging tools to perform therapies, including a VE, the Vybe haptic gaming pad [11], and the LMC gesture sensor [10], and (iv) provides a secure way for patients to replay their massage according to the therapy program.

3. Description of the System

Figure 1 shows the architecture of the real-time haptic-enhanced telerehabilitation system. The GoodVybesConnect system was designed following therapists' recommendations, which were obtained through a contextual study performed in a rehabilitation center [9]. The system is composed of a web-based application (Figure 2) and a local VE (Figure 3).

The web-based application is a therapy administration system in which therapists can manage their therapy sessions' schedules and their patients as well as their corresponding clinical records. In addition, therapists can configure the therapy by adjusting the visual feedback, intensity, and kind of music, among others. The web application also shows the results of patients' past therapy sessions through tables and graphs (Figure 2). Moreover, patients are able to schedule a therapy session, view their therapy record, access past therapy sessions, and communicate via chat with their therapist.

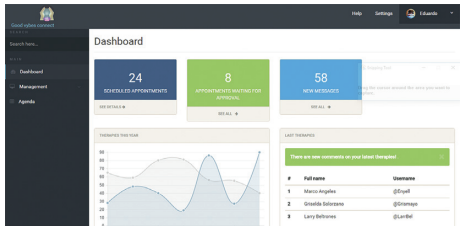


Figure 2: Therapy management web application. The therapist can accept, communicate, and update the patient's therapy record. In addition, the patient can request a therapy session, communicate with the therapist, and obtain his/her therapy files.

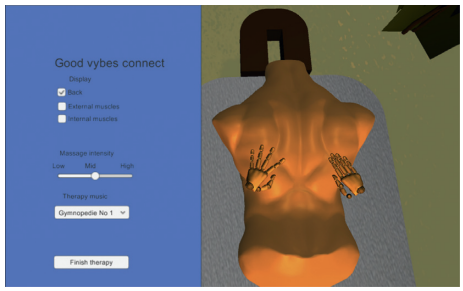


Figure 3: VE for the configuration and execution of back massage therapy.

The VE is a haptic-enhanced telerehabilitation system that was developed using Unity 5.3.4, the Vybe haptic gaming pad [11], and the Orion library for the LMC gesture sensor [10]. The main objective of the VE is to provide a real-time and remote back massage therapy to patients. The LMC gesture sensor is used as the input device, while the Vybe haptic gaming pad is used as the massage bed (output device) to execute therapy (Figure 1). The Vybe haptic gaming pad is a vibro-tactile grid display that has six voice-coil actuators located on the upper back and six DC motors located on the seat and lower back. The voice coils play smooth vibrations (150–250 Hz), while the DC motors play rumble-like effects [11].

Figure 3 shows the configuration screen in which the therapist chooses the virtual back that she/he wants to visualize (i.e., back muscles and surface or intermediate muscles), the intensity of the haptic feedback, and the music the patient will hear during therapy. The movements of the therapist's hands are detected with the LMC gesture sensor and transmitted online in real time to both the VE and the Vybe haptic gaming pad. All parameters can be changed online during execution of the therapy. After concluding the massage, the therapist stores the session and shares it through the web application with the patient so that she/he can reproduce his/her therapy again if indicated by his/her therapy program.

4. User Experience Evaluation with Potential Patients

4.1. Procedure. To obtain the participants' perception of usability and the participant's user experience when receiving a massage of the back with our system, we conducted a controlled usability test at the research laboratory of a local public university [9]. We used both questionnaires and controlled observation to evaluate whether the participant's user experience was pleasurable or not. That is, participants have a pleasant user experience when they feel relaxed and/or happy during the massage session, and they have an unpleasant user experience if they feel stressed, scared, or nervous.

Figure 4: Participant during back massage therapy. Materials: (1) video camera, (2) computer running the VE for back massage therapy, and (3) haptic Vybe gaming pad used as massage bed.

The potential patients were 25 older adults (age: mean ± SD: 63.68 ± 7.98 years, 11 males and 14 females), from whom 32% have received physical therapy due to various reasons (e.g., neck or spine problems, Parkinson's disease, or depression), 60% were under medical treatment, and 72% use the computer on a regular basis. All participants signed a consent form and agreed to be video-recorded during the execution of the back massage (Figure 4).

System Usability Scale (SUS) [27] and Technology Acceptance Model (TAM) [28] questionnaires (5-point Likert scale) were applied at the end of the study. During both the on-entry demographic questionnaire and the on-exit questionnaire, participants were asked about their emotional state at that moment, for example, if they were stressed, scared, nervous, relaxed, happy, or neutral. Besides, while answering the on-exit questionnaire, participants were asked for which of the aforementioned moods prevailed during the massage as well as their perception of the haptic feedback received through the Vybe haptic gaming pad. After that, two expert evaluators realized a controlled observation evaluation analyzing the video recordings to further verify the participants' moods.

4.2. Preliminary Results. From the on-exit questionnaire, we obtained the following results: 8% participants reported feeling neutral, 84% relax, and 8% happy. Further, during the analysis of the video recordings of the massage sessions, the expert evaluators identified that participants depicted the following moods: 92% happy, 96% relaxed, and 12% nervous. However, participant's moods were not constant throughout the entire massage sessions, as they were observed during varying time periods and sometimes in a combined form, whereby the experts selected which mood predominated during the massage session, determining thus that the participants were observed: 92% relaxed and 8% nervous. Table 1 shows the moods reported by participants and observed predominantly in them by the experts.

Table 1: The predominant moods reported by the subjects and observed by the experts (i.e., 1 = stressed, 2 = scared, 3 = nervous, 4 = neutral, 5 = relaxed, and 6 = happy).

Result	Participant																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Reported by the subject	4	5	5	5	5	5	6	5	5	5	5	5	6	5	5	5	5	5	5	5	5	5	4	5	5
Observed by evaluators	5	5	5	5	3	3	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5

According to Table 1, it can be observed that there was a coincidence of 80% (20 cases) between how the participants self-perceived and what was observed by the experts. From these 20 instances, 19 cases coincide that they were relaxed and were thus observed by the experts (5–5), while on the remaining cases (that of participant P23), it is reported as having a neutral state during the massage, which was ratified by the observers (4–4). In the first case, it is concluded that the user experience was pleasant since it was declared so by the participants and ratified by the experts; in the second case, in the absence of enough evidence, the user experience of the participant was declared as not pleasant.

Additionally, the differences found in the rest of the participants include the case of participant P1 (4–5), the cases of participants P5 and P6 (5–3), and the cases of participants P7 and P13 (6–5). For the case of participant P1 (4–5), neutral self-perception versus observed relaxed, at the beginning of the session, the participant reported that, on one side of the device, she/he did not feel the same way as on the other; however, after a minute, the participant closed his/her eyes and he/she seemed relaxed during the rest of the session; therefore, the observers reported him/her as relaxed; however, it is likely that the problem of the device persisted and that the experience was not pleasant enough as to become relaxed or feel happy; so, the participant was self-perceived as neutral and we concluded that the participant did not have a pleasant experience.

Of the two participants in the tuple (5–3), self-perceived relaxed versus observed nervous, observers reported that, throughout the session, participants were observed unease; in the case of P5, she was observed during the whole session with her eyes open wide and blinking fast, and at no time, she seemed relaxed; with P6, something similar occurred, as he also constantly squeezed the lips and moved the fingers of his hands, so he was clearly nervous. This nervousness could be caused by the expectation of not knowing how the device worked. Neither P5 nor P6 had received some type of treatment or physical therapy in the past. In both cases, we concluded that the user experience was not pleasant.

Finally, two participants reported feeling happy during the session; however, according to the observation, they were perceived very relaxed (6–5). It can be said that they were enjoying the massage a lot, and although they were relaxed, they perceived themselves happier than relaxed. In both cases, it is considered then that their user experience was pleasant.

As conclusion, we have that 84% of the participants had a pleasant user experience when reporting themselves as relaxed or happy during the massage session and that, for the 16% remaining participants, there is no evidence that they had a pleasant user experience by presenting clear evidence of nervousness throughout the session or reporting a neutral emotional state.

5. Evaluation with Specialists

5.1. Procedure. The study was conducted with specialists in motor rehabilitation who participated in a system-extended usability test using the on-exit questionnaire inspection technique. The purpose of this evaluation was to obtain the specialists' perception about the usability of the system and their user experience while participating both as patients and therapists. In this way, we aim at assessing the usability of the current prototype and identifying some insights to improve the system.

The participants were 10 female specialists (7 therapists and 3 psychologists, mean age: 29.9 ± 8.00 years) who work in two different institutions (7.25 ± 5.95 years average experience). The specialists provide motor or sensory therapy to various patients due to multiple etiologies (e.g., sensory disintegration, injury due to physical activity, lack of knowledge of the body, feeling exhausted, stress, and hypertension, among others). Seventy percent of them care for patients of all ages, while the rest provide care only for children with autism, social integration problems, and Down's syndrome, among others. Eighty percent of the specialists use common instruments to carry out the therapy, for example, swings, hammocks, textures, mattresses, and balls using their hands or knees, and the rest uses videos, songs, and stories to motivate patients. Finally, 100% consider that computer systems can be used as a support tool in therapy according to the diagnosis and treatment of the patient.

All participants signed a consent form and agreed to be video-taped during the execution of the evaluation. The evaluation was conducted at the work center of each specialist. The VE application was installed on a computer, while the Vybe haptic device was used as a massage bed (Figure 5).

In this evaluation, the specialists participated in two conditions: (a) C1, as patients and (b) C2, as therapists, so that they could have real experiences from both perspectives and enrich their perception of the system. Our hypotheses of the use of the system are as follows: (H1) the usability perception of the GoodVybesConnect system is higher for condition C1 than for condition C2 and (H2) the user experience of the participant in condition C1 is more pleasant than in condition C2.

The evaluation procedure consisted of a welcome message, initial explanation, and a questionnaire related to

Figure 5: Example of the assessment scenario with specialists. Materials: (1) video camera, (2) computer running the VE for back massage therapy, and (3) haptic Vybe gaming pad used as massage bed.

(a)

(b)

Figure 6: Specialists during back massage therapy. Scenario: (1) participant P9 as patient (C1) and participant P10 as therapist (C2) and (2) participant P10 as patient (C1) and participant P9 as therapist (C2).

their demographic data. Then, in pairs, the specialists used the system for 4 minutes as a patient (C1) and for 4 minutes as a therapist (C2) (Figure 6). Role assignment (i.e., therapist or patient) was randomly controlled to avoid bias. The intensity of the haptic feedback of the provided massage was also set randomly (i.e., ¼, ½, ¾, and 1). All participants experienced the same auditory feedback during the massage.

System Usability Scale (SUS) and Technology Acceptance Model (TAM) questionnaires (5-point Likert scale) were applied at the end of the study. During both the on-entry demographic questionnaire and the on-exit questionnaire, participants were asked about their emotional state at that moment, for example, if they were stressed, scared, nervous, relaxed, happy, or neutral. Besides, while answering the on-exit questionnaire, participants were asked (using a 5-point Likert scale) which of the aforementioned moods prevailed during the massage (as therapist and as patient) and their perception of the haptic feedback received through the Vybe haptic gaming pad. Additionally, participants were asked about the system functions and haptic device characteristics and comments or suggestions about the system. Finally, a primary analysis of normality (using an input analyzer tool) indicated that the outcome variables were not normally distributed. Therefore, to analyze the hypotheses, a nonparametric statistical analysis was applied using the two-tailed nonparametric Wilcoxon test [29].

Table 2: Results of the evaluation in terms of the usability perception (H1).

Hypothesis	Outcome H1 (usability)	
Condition	C1	C2
M	4.20	4.30
IQR3	4.75	5.00
IQR1	4.00	3.50

C1: specialist as patient; C2: specialist as therapist; M: median; IQR: interquartile range.

Table 3: Summary of Wilcoxon test on H1 hypothesis (critical value $T = 3$ with $N = 8$).

Comparative	Statistical value	Significance of the difference
H1 (usability, alpha = 0.05)		
C1 versus C2	14.5	Not significant

C1: specialist as patient; C2: specialist as therapist.

6. Results

6.1. Perception of Usability. Table 2 presents the summary of the participant's perception of usability regarding the system when using it as a patient (C1) and using it as a therapist (C2). As seen in Table 2, the usability perception is high in both conditions: as a patient (median 4.20/5) and as a therapist (median 4.30/5). In addition, the statistic test indicates that the differences in perception are not significant (see H1 in Table 3); therefore, H1 is rejected. Additionally, Figure 7 shows that although the differences in the usability perception of the GoodVybesConnect system were not significant, 70% of the specialists completely agree that the system is usable as a support tool in massage therapy, unlike their perspective as patients, where only 20% completely agree that the system is usable, and 60% of them agree.

All participants as patient (C1) stated that if they would have the system available, they would use it: 20% stated that they would use it once or twice a week, 30% stated that they would use it three or four times a week, 30% stated that they

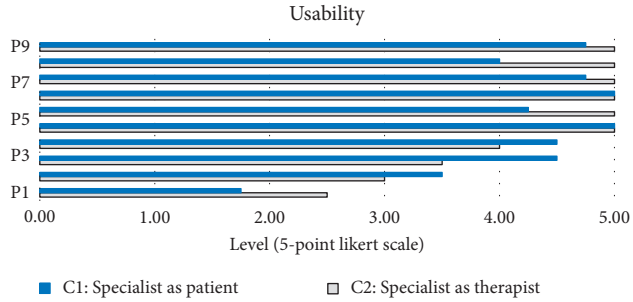


Figure 7: Usability perception (H1). C1: specialist as patient; C2: specialist as therapist. Level: 1 = totally disagree, 2 = disagreement, 3 = neutral, 4 = agree, and 5 = completely agree.

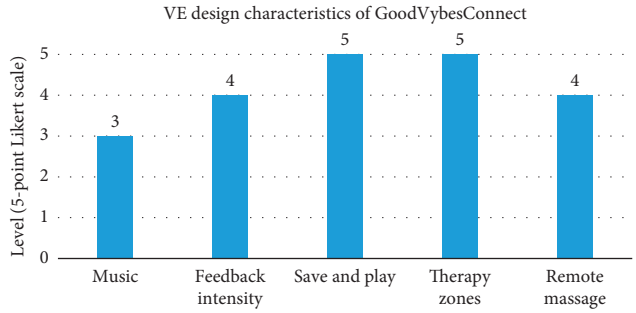


Figure 8: Importance level of design characteristics of the VE of GoodVybesConnect according to the specialist. Level: 1 = totally disagree, 2 = disagreement, 3 = neutral, 4 = agree, and 5 = completely agree.

would use it 5 times a week, and 20% stated that they would use it 6 times a week or more. Regarding time of use, 50% declared that they would use it for 10 minutes, 40% stated that they would use it for 15 minutes, and 10% stated that they would use it for 15 minutes or more.

Furthermore, 100% of the participants as therapist (C2) stated that if they would have the system available, they would use it: 10% stated that they would use it once or twice a week, 40% stated that they would use it three or four times a week, 40% stated that they would use it 5 times a week, and 10% stated that they would use it 6 times a week or more. Regarding time of use, 10% declared that they would use it for 5 minutes, 60% stated that they would use it for 10 minutes, and 30% stated that they would use it for 15 minutes or more.

In addition, regarding the use of the VE, the specialists indicated that they completely agree that the most important characteristics of the VE are saving and reproducing the therapy (median 5/5), as well as focusing on certain areas of therapy according to a therapy program (5/5). The specialists also agreed that it is important that the massage can be performed remotely (median 4/5) and that the levels of intensity of the haptic feedback can be adjusted (median 4/5). Finally, changing the music in real time was the characteristic perceived as less important (median 3/5) (Figure 8).

Regarding the perception of the haptic feedback of the system according to the specialists (Figure 9), a good perception of the haptic feedback was obtained (median 4.18/5).

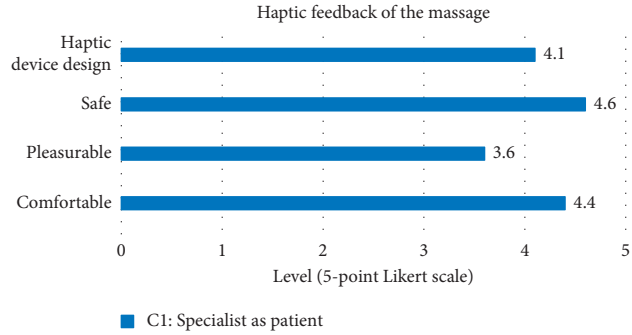


Figure 9: Massage haptic feedback perception (H2). C1: specialist as patient. Level: 1 = totally disagree, 2 = disagreement, 3 = neutral, 4 = agree, 5 = completely agree.

Figure 9 shows the average scores perceived by the specialists. According to these results, the specialists felt first and foremost the *safe* feature (median 4.6/5); second, they liked the *comfort* provided by the mattress (median 4.4/5); third, they liked the *design of the haptic device*; and finally, the *pleasurable* feature had the lowest perception (median 3.6/5).

Finally, several specialists indicated that systems such as the GoodVybesConnect could be very useful to attend children who are unaware of their own body, who do not let themselves to be touched, or who have suffered any injuries, among others. In addition, in adult patients, it can be used to treat, for example, stress, anxiety, tension, muscle pain, or to relax patients who are prostrate due to illness.

6.2. User Experience. Table 4 shows the summary of the specialists' user experience upon receiving massage (condition C1) and providing massage (condition C2) (H2 hypothesis). In addition, Table 5 presents the result of the analysis in terms of the emotions presented by the specialists. As shown in Table 5, the results obtained show evidence that the differences between emotions during therapy on receiving (C1) or providing (C2) therapy are not significant, and therefore, H2 is discarded. However, on average, the user experience in condition C1 (receiving the massage) was slightly more pleasurable than in condition C2 (providing the massage).

Figure 10 shows the emotions of the specialists when using the GoodVybesConnect (H2) system. On the one hand, while feeling the mattress with haptic feedback and receiving the massage (C1), everyone denied feeling stressed (100%), and most participants denied having felt fear (90%) or being nervous (80%). Likewise, it can be emphasized that most of them affirmed to have felt relaxed, (90%) and all affirmed to have felt happy (100%). On the other hand, when using the system and providing the massage (C2), most of them denied having felt stressed (90%) or having felt fear (80%); while some reported feeling nervous (30%) or neutral (60%). Finally, regarding feeling relaxed, 30% disagreed and 40% agreed or completely agreed; while most of them (70%) indicated that they felt happy.

Table 4: Results of the evaluation in terms of the emotions that the participants felt during the use of the system (H2).

Hypothesis	Outcome H2 (user experience)	
Condition	C1	C2
M	2.58	2.57
IQR3	2.67	3.00
IQR1	2.33	2.33

C1: specialist as patient; C2: specialist as therapist; M: median; IQR: interquartile range.

Table 5: Summary of Wilcoxon test on H2 hypothesis (critical value $T = 0$ with $N = 6$).

Comparative	Statistical value	Significance of the difference
H2 (user experience, alpha = 0.05)		
C1 versus C2	7	Not significant

C1: specialist as patient; C2: specialist as therapist.

7. Discussion

7.1. Perception of Usability. The main results suggest that the usability perception of the GoodVybesConnect system is high, both when using it as a patient (C1) (median 4.20/5) and as a therapist (C2) (median 4.30/5). Sixty percent of the specialists while using the system to provide the massage (condition C2) were in complete agreement (i.e., median 5/5) that the system is useful, easy to use, easy to learn, and with high intention to use (Figure 7). No significant differences were found regarding the usability perception of the GoodVybesConnect system in both conditions (C1 and C2), since in both cases the perception of usefulness was high, and given that all the specialists indicated that they were willing to use the system once it becomes available.

In addition, the specialists evaluated the design features of the current system, indicating the following as the most important: (a) saving and reproducing a therapy session (median 5/5), (b) targeting therapy areas according to a program (Figure 8), (c) that the massage could be performed remotely (median 4/5), (d) that the intensity of the haptic feedback could be adjusted, and (e) that music could be changed in real time, with the latter being considered less important (median 3/5). As evidenced by these results, on average, the design characteristics of the system were considered important by the specialists.

The main findings in the perception of haptic feedback suggest that 50% of the specialists while receiving massage completely agree that the intensity is adequate (median 4.18/5) (Figure 9). However, the rest of the specialists indicated that it is necessary to increase and focus on the intensity of the massage feedback to generate more pleasant experiences of use. On the other hand, although the perception of the design of the device is adequate (median 4.3/5), they indicate that it could have larger dimensions to cover other massage zones, including the neck, sides of the trunk, nape, and legs.

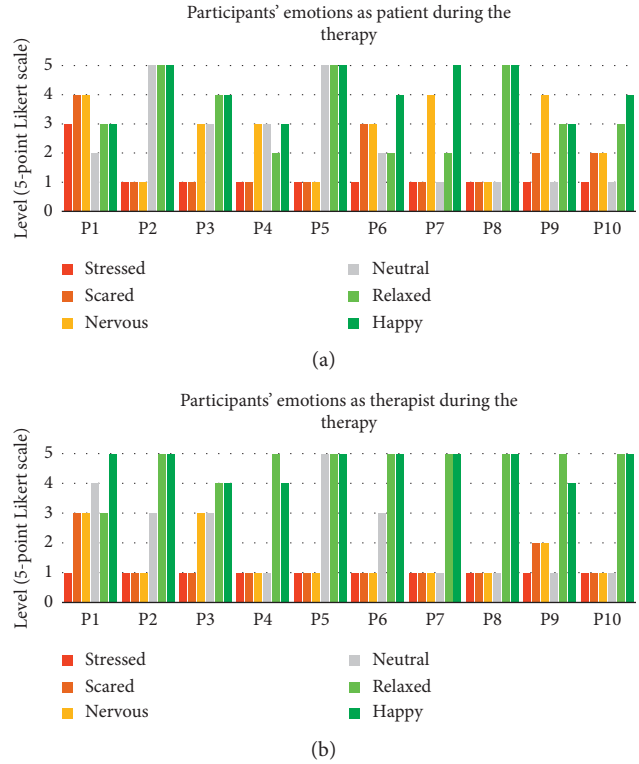


Figure 10: Emotions during therapy (H2). Left (C1: specialist as patient); right (C2: specialist as therapist). Level: 1 = totally disagree, 2 = disagreement, 3 = neutral, 4 = agree, 5 = completely agree.

7.2. User Experience. The analysis of the specialists' emotions upon receiving massage (condition C1) and on providing massage (condition C2) indicates that most specialists did not had a negative experience of use (i.e., stress or fear) and that their differences are not significant. However, P1 in his role of therapist (C2) expressed that he felt scared, which can be explained according to his final comments "I was afraid to do an incorrect movement when giving the massage" (P1). In addition, in the therapist role, the specialists P1, P7, and P9 answered that they were nervous when using the system, which could be explained based on some of their comments: "[it is necessary] to accompany the system with a user manual, or with a video tutorial, as not all therapists are used to interact with a virtual environment" (P7), "I found everything very good, it just made me nervous not to do it correctly" (P9) (Figure 10). This can also be explained by the fact that during the use of the GoodVybesConnect system, the specialists indicated feeling more relaxed when receiving the massage (88%) than when giving the massage (50%) (Figure 10). Additionally, in their perception while receiving the massage (C1), 90% felt relaxed and 100% felt happy. This higher percentage in positive experience of use might explain why at the end of the massage participants P2, P4, P5, P6, P7, P8, P9, and P10 asked to continue the massage; some of their comments were "why it ended" (P4, P5, P8, and P10), "I was falling asleep" (P2, P9, and P10), "it felt great, when will it be available?" (P6), and "where can I buy it?" (P5, P9, and P10). Finally, several of the specialists mentioned that they felt better when the intensity of the haptic feedback was

greater and indicated that it would be advisable to increase it to improve its pleasurable feature.

According to the results of both evaluations, there is evidence that, on the one hand, the GoodVybesConnect system was well perceived by potential users and that it generated mainly pleasant user experiences. On the other hand, the specialists agreed on its high usability from the two perspectives (C1 and C2), as well as on that it generated pleasant user experiences from the perspective of the patient. Specialists have also shared different proposals aiming at improving the design of the VE and of the massage cushion.

8. Design Insights

Based on the comments and suggestions made by specialists in rehabilitation, the following system design topics emerged:

(i) *Adjustment of therapy strategies*. The specialists suggested that it would be interesting if the patient could provide the massage in the absence of a therapist, or even have both the patient and the therapist able to select different preset therapy techniques to treat the patient. Some of the comments were: "it would be great if at any time there was the option of being able to direct your own [massage] when you do not have access to a therapist" (P6), "that the system could control itself to receive the massage" (P8), and "generate a prerecorded massage to use it with other patients" (P7).

(ii) *Selection, delimitation, and adjustment of the massage zones*. Specialists said that it is necessary to focus on the massage zones to avoid touching sensitive areas of the patient. They also suggested that it would be important to add the area of the head, neck, nape, trunk sides, and legs. They also recommend that the delimitation of the zones can be determined depending upon the corpulence of the person and their age, for example, child or adult. Some of the comments were: "inquire about the affected areas to give more emphasis" (P1), "isolate the stimulus of the pressure being exerted" (P2), "exert less intensity in the central zone (column)" (P7), and "consider the scapula" (P8).

(iii) *Adequate calibration and size of the massage device*. Specialists suggested that the haptic device be wider and longer. In addition, they mentioned that, once the device is larger, different dimensions can be configured to adapt the treatment according to the patient and add [supports for] the neck, sides of the trunk, shoulders, or legs. Some of the comments were: "I would like to add an attachment for the neck" (P2), "it is necessary to think about overweight people (obesity) because [in its current version] it may not cover their expectation for treatment" (P3), "the dimensions are suitable for standard people" (P3), "[it requires a] wider and longer mattress" (P4, P5), "maybe add [support] for sides and shoulders" (P6), "increase its dimensions [haptic device]" (P10), "the mattress could be wider for the benefit of larger people" (P10), and "[it] could cover the area of the legs" (P10).

(iv) *Add temperature and increase the intensity of the haptic feedback of the massage*. Specialists suggest that the stimulus for the patient was better at the higher intensity of the massage; in addition, they recommended adding temperature to strengthen the relaxation of the patients. Also, they considered that it is necessary to numerically quantify the intensity of the haptic feedback. Some of the comments were: "[it requires] intensity levels of feedback, consider them as a wider numeric scale rather than low, medium, and high" (P3), "higher feedback intensity of the mattress" (P1, P2, and P8), "I would like to add temperature to strengthen relaxation" (P2), "I felt better when the intensity of the massage was at the highest" (P3, P5, and P9), and "to relax more in the upper part of the back" (P8).

(v) *Add haptic feedback to provide support to the movements of the therapist and improve the visual representation in the VE*. The fear and nervousness emotions of the specialists while using the VE of GoodVybesConnect suggest that it is necessary to integrate the haptic feedback while providing the massage so that the therapist becomes aware of the intensity with which she/he is touching the patient. In addition, this intensity of the haptic feedback must correspond to the visual representation of the patient's virtual body in the VE. Some of the comments were: "I was afraid to do the movement incorrectly when giving the massage" (P1) and "I found everything very good, I just felt nervous of not doing it correctly" (P9).

9. Conclusions

A real-time haptic-enhanced telerehabilitation system for massage therapy was evaluated with potential users and specialists. In the evaluation with potential users, the experience of use was obtained, while with the specialists in addition to the experience of use, the perception of usability of the system was determined, along with some insights to improve its design.

To the best of our knowledge, GoodVybesConnect is the first system that allows therapists to perform back massage therapy in real time and remotely. Although we have evaluated the system with a small group sample, namely, 25 potential patients and 10 specialists, we would like to point out that, in the clinical literature domain, there are works where similar statistical tests are applied with a similar number of participants, for instance [5, 17, 19]. In addition, the HCI literature suggests that the inclusion of five experts during the evaluation stage makes it possible to detect 80 to 85% of the usability problems of a design proposal [30].

Furthermore, the results of both evaluations provide evidence of a high perception of usability, positive user experience (i.e., relaxation and joy), and good haptic feedback (i.e., safe, pleasurable, and comfortable) provided by

the system. In addition, the comments suggest that the specialists perceived the current design features as important, and they generated new design insights focused primarily on improving massage feedback. Finally, as future work, it is necessary to further include the emerged design insights to enhance the current system. First, we will strengthen the design of the system to improve the experience of use and its clinical impact. Second, we are going to evaluate the proposed tool with more specialists (i.e., clinicians and therapists), and potential patients, to confirm the observed trends. Third, we will use the system for a longer period of time to better establish the scope and impact of these results.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors thank the therapists from the two rehabilitation centers at Ensenada, México, for participating in the study. This work was funded by Universidad Autónoma de Baja California under Grant 212 Proyectos de Servicio Social 2016 and CONACYT scholarship number 344565 provided to the first author.

References

[1] M. A. McCaskey, C. Schuster-Amft, B. Wirth, Z. Suica, and E. D. de Bruin, "Effects of proprioceptive exercises on pain and function in chronic neck- and low back pain rehabilitation: a systematic literature review," *BMC Musculoskeletal Disorders*, vol. 15, no. 1, p. 382, 2014.

[2] M. Rogante, M. Grigioni, D. Cordella, and C. Giacomozzi, "Ten years of telerehabilitation: a literature overview of technologies and clinical applications," *NeuroRehabilitation*, vol. 27, no. 4, pp. 287–304, 2010.

[3] D. C. Cherkin, K. J. Sherman, J. Kahn et al., "Effectiveness of focused structural massage and relaxation massage for chronic low back pain: protocol for a randomized controlled trial," *Trials*, vol. 10, no. 1, p. 96, 2009.

[4] M. Romanowski, J. Romanowska, and M. Grześkowiak, "A comparison of the effects of deep tissue massage and therapeutic massage on chronic low back pain," *Studies in Health Technology and Informatics*, vol. 176, pp. 411–414, 2012.

[5] D. F. Zullino, S. Krenz, E. Frésard, E. Cancela, and Y. Khazaal, "Local back massage with an automated massage chair: general muscle and psychophysiologic relaxing properties," *Journal of Alternative and Complementary Medicine*, vol. 11, no. 6, pp. 1103–1106, 2005.

[6] P. Buselli, R. Bosoni, G. Busè et al., "Effectiveness evaluation of an integrated automatic thermomechanic massage system (SMATH® system) in non-specific sub-acute and chronic low back pain - a randomized double-blinded controlled trial, comparing SMATH therapy versus sham therapy: study protocol for a randomized controlled trial," *Trials*, vol. 12, p. 216, 2011.

[7] D. J. Engen, D. L. Wahner-Roedler, A. M. Nadolny et al., "The effect of chair massage on muscular discomfort in cardiac sonographers: a pilot study," *BMC Complementary and Alternative Medicine*, vol. 10, no. 1, p. 50, 2010.

[8] S. McEwen, "Social work in health care when conventional meets complementary: nonspecific back pain and massage therapy," *Health and Social Work*, vol. 40, no. 1, pp. 19–25, 2015.

[9] C. Ramírez-Fernández, E. García-Canseco, A. L. Morán et al., "GoodVybesConnect: a real-time haptic enhanced tele-rehabilitation system for massage therapy," in *Proceedings of the 10th International Conference, UCAmI 2016*, pp. 487–496, San Bartolomé de Tirajana, Gran Canaria, Spain, November–December 2016.

[10] A. H. Smeragliuolo, N. J. Hill, L. Disla, and D. Putrino, "Validation of the Leap Motion Controller using markered motion capture technology," *Journal of Biomechanics*, vol. 49, no. 9, pp. 1742–1750, 2016.

[11] Disney, "Disney research," 2015, https://www.disneyresearch.com/.

[12] W. L. Chen, G. J. Liu, S. H. Yeh, M. C. Chiang, M. Y. Fu, and Y. K. Hsieh, "Effect of back massage intervention on anxiety, comfort, and physiologic responses in patients with congestive heart failure," *Journal of Alternative and Complementary Medicine*, vol. 19, no. 5, pp. 464–470, 2013.

[13] K. G. Davis and S. E. Kotowski, "Preliminary evidence of the short-term effectiveness of alternative treatments for low back pain," *Technology and Health Care*, vol. 13, no. 6, pp. 453–462, 2005.

[14] J. Baranowsky, P. Klose, F. Musial, W. Haeuser, G. Dobos, and J. Langhorst, "Qualitative systemic review of randomized controlled trials on complementary and alternative medicine treatments in fibromyalgia," *Rheumatology International*, vol. 30, no. 1, pp. 1–21, 2009.

[15] L. Gaskell, S. Enright, and S. Tyson, "The effects of a back rehabilitation programme for patients with chronic low back pain," *Journal of Evaluation in Clinical Practice*, vol. 13, no. 5, pp. 795–800, 2007.

[16] Y. Noto, M. Kudo, and K. Hirota, "Back massage therapy promotes psychological relaxation and an increase in salivary chromogranin A release," *Journal of Anesthesia*, vol. 24, no. 6, pp. 955–958, 2010.

[17] H. Ishii, H. Koga, Y. Obokawa, J. Solis, A. Takanishi, and A. Katsumata, "Path generator control system and virtual compliance calculator for maxillofacial massage robots," *International Journal of Computer Assisted Radiology and Surgery*, vol. 5, no. 1, pp. 77–84, 2010.

[18] J. Wang and Y. Li, "Massaging human feet by a redundant manipulator equipped with a tactile sensor," in *Proceedings of the 2010 IEEE/ASME International Conference*, pp. 7–12, Montreal, Canada, July 2010.

[19] R. C. Luo, C. W. Hsu, and S. Y. Chen, "Electroencephalogram signal analysis as basis for effective evaluation of robotic therapeutic massage," in *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2940–2945, Daejeon, Republic of Korea, October 2016.

[20] H. Gao, S. Lu, G. Tian, and J. Tan, "Vision-integrated physiotherapy service robot using cooperating two arms," *International Journal on Smart Sensing and Intelligent Systems*, vol. 7, no. 3, pp. 1024–1043, 2014.

[21] C. S. Meera, P. S. Sairam, and M. K. Gupta, "Path planning and motion control for a 3 DOF massaging robot," in *Proceedings of the 2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA)*, pp. 1–6, Amritapuri, Kerala, India, December 2016.

[22] T. Teramae, D. Kushida, F. Takemori, and A. Kitamura, "Estimation of comfortable/uncomfortable feeling based on

EEG by using NN and k-means algorithm for massage chair," *IEEJ Transactions on Electronics, Information and Systems*, vol. 131, no. 1, pp. 102–108, 2011.

[23] H. Jaafar, A. Fariz, S. A. Ahmad, and N. A. M. Yunus, "Intelligent massage chair based on blood pressure and heart rate," in *Proceedings of the 2012 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 514–518, Langkawi, Malaysia, December 2012.

[24] J. Kim, K. B. Lee, S. Lee, H. Yang, and S. G. Hong, "A novel stress measurement system with handhold electrodes in massage chairs," in *Proceedings of the 2016 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 859–863, Jeju Island, Republic of Korea, October 2016.

[25] R. C. Luo and C. C. Chang, "Electromyographic signal integrated robot hand control for massage therapy applications," in *Proceedings of IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems (IROS 2010)*, pp. 3881–3886, Taipei, Taiwan, 2010.

[26] C. Clar, A. Tsertsvadze, R. Court, G. Hundt, A. Clarke, and P. Sutcliffe, "Clinical effectiveness of manual therapy for the management of musculoskeletal and non-musculoskeletal conditions: systematic review and update of UK evidence report," *Chiropractic and Manual Therapies*, vol. 22, no. 1, p. 12, 2014.

[27] J. Brooke, "SUS-a quick and dirty usability scale," *Usability Evaluation in Industry*, vol. 189, no. 194, pp. 4–7, 1996.

[28] F. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.

[29] J. H. Zar, *Biostatistical Analysis*, Prentice-Hall, Upper Saddle River, NJ, USA, 2010.

[30] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 249–256, Seattle, WA, USA, March 1990.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 3185481, 15 pages
https://doi.org/10.1155/2017/3185481

Research Article

Semantic and Virtual Reality-Enhanced Configuration of Domestic Environments: The Smart Home Simulator

Daniele Spoladore,[1] Sara Arlati,[1,2] and Marco Sacco[1]

[1]*Institute of Industrial Technologies and Automation, Italian National Research Council, Via Corti 12, 20133 Milano, Italy*
[2]*Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy*

Correspondence should be addressed to Daniele Spoladore; daniele.spoladore@itia.cnr.it

Received 28 July 2017; Accepted 29 October 2017; Published 19 December 2017

Academic Editor: Pino Caballero-Gil

This paper introduces the Smart Home Simulator, one of the main outcomes of the D4All project. This application takes into account the variety of issues involved in the development of Ambient Assisted Living (AAL) solutions, such as the peculiarity of each end-users, appliances, and technologies with their deployment and data-sharing issues. The Smart Home Simulator—a mixed reality application able to support the configuration and customization of domestic environments in AAL systems—leverages on integration capabilities of Semantic Web technologies and the possibility to model relevant knowledge (about both the dwellers and the domestic environment) into formal models. It also exploits Virtual Reality technologies as an efficient means to simplify the configuration of customized AAL environments. The application and the underlying framework will be validated through two different use cases, each one foreseeing the customized configuration of a domestic environment for specific segments of users.

1. Introduction

Ambient Assisted Living (AAL) is a research field emerged in the 1990s and acquiring growing importance. It proposes residential tools and solutions devoted to the improvement of everyday life, focusing on the person and his/her interactions with both technologies and domestic environment. In this context, AAL solutions aim at enhancing the dwellers' quality of life, comfort, and well-being and can be customized to address specific problems for particular segments of population, such as elderlies or people affected by disabilities. However, the development of AAL systems is not an easy task, since it has to take into account some fundamental features. The first regards the end-users for whom the systems are being developed: each solution, in fact, must be tailored on the dwellers and their needs. It is therefore imperative to adopt a proper paradigm able to be accountable for the users' real needs and capabilities. Another relevant issue gathering the attention of many researchers concerns the possibility to get full interoperability among the devices deployed in the house. Furthermore, a third issue arises when it comes to validating the abovementioned interoperability: the costs and time to set up a domestic environment are relevant and may result dissuasive.

In order to tackle these three major issues, the D4All project relied on a consolidated framework (Section 3.5), which leverages the integration capabilities of Semantic Web technologies and the possibility to model relevant knowledge (about both the dwellers and the domestic environment) into formal models. Furthermore, the project framework took into account the possibilities offered by Mixed and Virtual Reality technologies to develop an application dedicated to the configuration and customization of domestic environments—the Smart Home Simulator.

The remainder of the paper is structured as follows: Section 2 offers a review of the use of Semantic Web technologies and Virtual Reality for the configuration of AAL systems. Section 3 delves into the aims, issues, solutions, and architecture addressed by the D4All project. Section 4 describes in detail the semantic knowledge base, backbone of the D4All project, while Section 5 depicts the role of the Virtual Reality. Section 6 proposes real use cases developed to test the efficiency of the Smart Home Simulator and its framework. Finally, the conclusions summarize the main outcomes and sketch the future works that will be pursued.

2. Related Works

This section highlights the most relevant works in the field of semantic model for the description of dwellers and users of AAL solutions, interoperability among appliances, AAL systems, and the use of Virtual Reality technologies for the simulation and configuration of living environments and smart homes.

Semantic Web technologies have proved to be an efficient way to represent the knowledge related to the domain of persons and houses, and they have been exploited in a variety of works. In particular, the semantic modelling of the smart home's dwellers and their activities has covered an increasing importance in the last decade. In [1], the authors focused on the detection of human activity inside a home, representing them into a set of ontologies covering the contexts of the smart home and activities of daily living. This model allows to infer, among other features, activity-concerned information, thanks to reasoning processes over raw data and contextual information. Razmerita et al. [2] developed OntobUM, an ontology-based architecture providing information on the user's identity, contacts, preferences, and competencies; this model is partially generated by the user and several intelligent services, which automatically update the information about the user taking into account his/her behaviour. In [3], an ontology for the modelling of AAL services is presented; this ontology considers concepts necessary to describe the environments, the users involved, and their activities in the environment, as well as their habits and abilities. A focus on the modelling of the user's static information, such as physical characteristics, living condition, profession, interests, education, and user experience, is provided in [4]. This work provided a methodology for an ontology-based activity of user profiling able to avoid omissions and errors. Skillen et al. [5] proposed an ontology for the personalization of context-aware applications, which considered both static and dynamic aspects of the users. The model takes into account users' health, interests, preferences, and abilities. In the field of context-aware adaptive systems, in [6], it is described a user-profile ontology able to represent situation-dependent sub-profiles; the ontology allows to automatically trigger the personalization of services according to the users.

Also in the field of interoperability among devices, the possibility to represent domain knowledge offers several advantages, described in various works. Welge et al. [7] addressed the issue of interoperability of different devices and distributed systems identifying in the exploitation of knowledge management and ontologies the key to overcome this problem. Semantic Web technologies are presented in [8] as an effective approach to allow the management of heterogeneous information in Ambient Intelligence solutions; in this work, the FLERSA tool [9] is enhanced in its interoperability features, allowing the deployment of Ambient Intelligence applications making use of different technologies and platforms. In [10], the authors described an upper ontology for the semantic support of AAL platforms, capturing the semantics of the AAL domain and depicting the services provided by a system and including the orchestration of data coming from sensors and other services. In [11], a result of the European COMANCHE project, an ontology-enabled knowledge base allows the modelling of a domestic environment with detail on the services provided to its inhabitants. The knowledge base allows the conflict-free and up-to-date functioning of the home network while providing a description of the appliances and the relationships among them.

Dealing with the configuration and the validation of Smart Homes, in the last decade, researchers have developed several systems with the aim of testing their solutions before implementing them in the real world. Anticipating the validation of the designed systems, comprising different smart objects and sensors exchanging data in real time, in fact, reduces time and costs of a solution otherwise potentially very onerous. In [12, 13], the authors developed two context-aware simulators (ISS and CAAS) that are able to handle concurrently data coming from the sensors and home inhabitants to respond to specific user's needs; they were both able to respond to specific sensors' measurement also detecting and handling potential conflicts of operating rules, which may arise when conflictual actions are triggered by multiple input data. Although these two solutions were provided with a graphical-user interface and a simplistic 2D representation of the house, they do not make use of virtual reality. Instead, Sernani et al. developed a VR-based system to validate their expert system, named, Virtual Carer, which manages a distributed network of smart objects inside a smart home, where each component is modelled as an autonomous agent [14]. To recreate a realistic scenario, they also simulated the human behaviour taking into account a model based on human basic needs to trigger specific activities inside the smart home. Human behaviour was simulated also in [15] with the final aim of developing a low-cost system capable of synthesizing a dataset for activity recognition research in smart environments. In [16], the authors demonstrated that virtual environments can be a promising tool in the design phase, since they provide the stakeholders with a demonstration of the system functionalities, thus improving the final design of the solution—a nursing home, in this case—through the users' feedback and suggestions. In [17], the authors presented a Simulation Control Panel (SCP) to be integrated in the authoring tool in order to ease the validation of the designed solution by creating a direct link between the modelling and the testing environments. Finally, VR-based simulations were also employed to test the feasibility and the reliability of brain-computer interfaces in controlling smart home appliances and services with the final aim of restoring environmental control for subjects with severe disabilities [18, 19].

3. The D4All Project: Aims and Solutions

3.1. The D4All Project: Applying Universal Design to Domestic Environments. Traditional design of products and solutions is usually oriented toward standard individuals, which are abstraction of real men and women. This approach fails to consider the many variables regarding real end-users, such as their skills, knowledge, social interactions, and requirements. To overcome this limitation, the paradigm of

"Universal Design" [20]—also known as "Design For All"—has started to stand out: it aims at taking into proper consideration the different features characterizing the real human users, focusing on proposing solutions able to adjust to the specific needs of users. The main principles of this paradigm can be found also in the field of AAL, a discipline addressing the issue of increasing the quality of life of people in all stages of the lifecycle and providing them with assistive systems for an independent life, according to their abilities [21, 22]. Thus, AAL systems aim at finding efficient solutions to help elderly or impaired people to maintain an independent and autonomous living. Although many efforts have been made toward the development of AAL solutions, most of the systems and devices created are not able to take into consideration the real needs of their users and neglect the side of human interaction and real capabilities [23].

In this context, Design For All (D4All) project [24] aims at applying the guidelines of Universal Design into domotic and inclusive domestic environments. These environments, grouped under the term "Smart Home" [25], should be able to anticipate and respond to the needs of their dwellers, to promote their comfort and well-being during many activities of daily living (ADL) [26].

The design of a Smart Home for normally endowed people, families, elderlies, and people with impairments or disabilities requires a set of heterogeneous tools during its whole lifecycle process (e.g., concept, design, implementation, and test [17]). Furthermore, it requires paying great attention to the specific users who interact with the services and technologies made available, as well as they interact with these functionalities. Consequently, it is fundamental to manage in a coherent and efficient way the vast amount of diverse data that these interactions can generate, making the data available to both the end-users (the inhabitants) and the remote users (such as caregivers or clinical personnel). The following subsections delve into the specific issues belonging to these two branches, respectively, the proper identification of users' needs, with particular attention toward being physically active, and requirements and the possibility to make the appliances and tools of a domestic environment fully interoperable among each other.

3.2. Users' Needs and Requirements. Benefits of being physically active are widely known for people of all ages; different studies have demonstrated how physical activity can prevent the onset of several chronic illnesses such as hypertension, obesity, diabetes, osteoporosis, some forms of cancer, and cardiovascular diseases [27]. This is particularly true for elderlies, since ageing leads to the structural and functional deterioration of many physiological systems, even in absence of a specific pathology. This clinical condition is defined as *frailty*, a term that indicates a status of increased vulnerability, in which not only the risk of an adverse event (e.g., minor infection and falls) is increased but also the body response to a small insult results in a disproportionate change in the health condition (from independent to dependent or from lucid to cognitive impaired) [28]. In this context, it is indeed true that no intervention can stop the physical and cognitive decline related with age. However, there is strong evidence that a minimization of risks of chronic illnesses and disabilities can be obtained through regular exercise. Both the World Health Organization and the American College of Sport and Medicine provided recommendations for regular physical activity in elderlies. They underline its importance as preventive measure against physical and cognitive decline [29, 30], the loss of autonomy in daily living, and thus, the economic burden that the ageing of the world population has risen in recent years [31].

This concept holds also for healthy and people suffering from different chronic pathologies, although it is clear that for this last category of home dwellers, the provision of a supervised physical exercise may not be enough to guarantee their autonomy during ADL. In all these cases, the Smart Home must be enhanced with tailored solutions, addressing the coping with specific issues. For instance, for visually impaired users—if not completely blind—there is the need to light up efficiently every part of the house, especially dangerous areas such as steps. Deaf people need visual or haptic alarm to replace normal alerts [32]. Motor-impaired users who use a wheelchair require an *accessible* house, where spaces are designed to allow an easy manoeuvring of the wheelchair and where appliances and controls are reachable while staying seated [33].

3.3. Interoperability and Cooperation among the Appliances in the Smart Home. A Smart Home is expected to be a domestic residence equipped with a set of appliances—often called Smart Objects; as mentioned above, the Smart Home and its appliances work to ensure the dwellers' comfort and personalized living conditions. In traditional houses, the various appliances deployed are able to perform their tasks in a separate and isolated way. On the contrary, inside a Smart Home, the appliances are required to work together in a reliable and predictable behaviour [34] and to acquire, handle, communicate, and share the knowledge about the home inhabitants in order to meet the goal of achieving their comfort and well-being [35]. Therefore, appliances must be able to provide tailored services to improve the dwellers' assistance for a better, healthier, and safer life in their everyday living environments and must cooperate with several other appliances—possibly specialized and multivendor.

Nevertheless, this kind of synergy among distributed appliances is nowadays guaranteed only in domestic environments where appliances with the same communication pattern and protocol stack [36] have been deployed. The desired interoperability among appliances should comprise not only mere communication interoperability but also data and information models, as well as services provided and discovery mechanisms. Reaching this level of interoperability is not an easy task: it takes a lot of time and requires many design decisions to be made to accommodate the constrained nature of specific devices in a certain usage scenario.

3.4. The Project's Approach and the Smart Home Simulator. D4All faces the above challenges relying on a framework that encompasses the description of the users'

physiological status, the appliances, and the services they can provide. The goal of the framework (and of the application derived from it) is to provide users characterized by frailty or impairment with appliances and services able to cope with their impairments, thus helping them in performing activities that would otherwise be unattainable or ponderous. The solution is not limited to specific categories of users but extends its services to the whole spectrum of house dwellers. This goal is achieved through the configuration of the dwellers' house that allows to analyse both whether the specific user's physiological requirements are satisfied and the behaviour of the interoperable appliances before their hard deployment in the real house.

The need to properly assess the user's health condition in a qualitative and quantitative way was addressed by recurring to the International Classification of Functioning, Disability and Health (ICF), a holistic World Health Organization's framework providing a unified and standard language for the description of health-related components. ICF conceptualizes the functioning of an individual as a "dynamic interaction between a person's health condition, environmental factors, and personal factors" [37], acting as a tool able to ease the communication among the health stakeholders (caregivers and clinicians) and providing a standard and worldwide comparable description of the functional experiences of the individuals. Due to its vocabulary, which is easily interpretable also by nonclinical personnel, the classification can also be used in various health-related domains, such as rehabilitation [38] or reintegration of injured workers in workplaces [39]. The classification is organized in two main parts: the first, "*Functioning and Disability*," provides a description of the components *Body functions*, *Body structures,* and *Activities and participation*; while the second, "*Contextual Factors*," provides the means to describe the impact of the components *Environmental factors* and *Personal factors*. Each component is further deepened into chapters, which identify the addressed domain. The functioning of a person is then described through the interaction between his/her health condition and the context where he/she acts. Each component is identified by a letter (*b* for Body functions, *s* for Body structures, *e* for Environmental factors, and *d* for Activities and participation) and can be detailed by adding digits (Figure 1).

According to the number of digits following the letter, it is possible to get a code, whose length indicates the level of granularity—up to five digits. The functioning or disability of an individual can be assessed selecting the suitable category and its corresponding code and then adding a qualifier (from 0, meaning "absence of impairment," to 4, indicating a "complete impairment").

In addition to the characterization of the users performed using ICF, the knowledge about their physical status was deepened taking into account the cardiorespiratory fitness (CRF). As mentioned in Section 3.2, in fact, the general physical condition plays an important role in preventing the occurrence of several pathologies and, thus, must be taken in consideration not only when describing the user but also when designing the Smart Home, which should be built to take care of the users' health and well-being too [40]. CRF

b «*Body function*»	Component
b2 «*Sensory functions and pain*»	Chapter - first level item
b210 «*Seeing functions*»	Second level item
b2102 «*Quality of vision*»	Third level item
b21020 «*Light sensitivity*»	Fourth level item

Figure 1: An example of the structure of ICF.

expresses the physical fitness of a person, specifying his/her ability to carry out a dynamic, moderate-to-high intensity exercise over a prolonged period of time [41]. In addition, it has been demonstrated to be inversely proportional to the risk of mortality and of cardiovascular, pulmonary, and coronary diseases [42]. Therefore, CRF represents a suitable indicator of a subject general health status and can be used to monitor the progress or the regression of the health status of the Smart Home inhabitants through time [41].

Semantic Web technologies [43], in particular the modelling of domain knowledge into ontologies [44], were chosen to provide a set of formal and sharable descriptions of the concepts and their relationships composing the domains of the domestic environment, the user, and its appliances. Semantic Web technologies can indeed provide a description of the functioning of the appliances and their services, thus enhancing the semantic interoperability among them. Furthermore, with the adoption of Semantic Web Rule Language (SWRL) rules, it is possible to trigger specific inferences, such as the deployment of specific appliances and services to address the user's particular needs (as further illustrated in Section 3). However, the interaction of the appliances distributed in a cooperative domestic environment cannot always be validated in a real environment—mainly because of the high costs and time to set the environment up. In order to ease an a-priori evaluation, as well as to validate the design of integrated appliance solutions providing AAL services to users and their services, the Smart Home Simulator (SHS) has been developed. The aim of this application is to allow designers of domestic environments to simulate and configure a house by taking advantage of a virtual representation of the house and the appliances. SHS also allows the designer to tailor the services offered by each appliance, leveraging the descriptions provided by Semantic Web technologies; in this way, the SHS allows the construction of a complete and clear picture of the status of the assisted users and their living environment before effectively hard-deploying the customized AAL solution. The result is a realistic simulation of a home where the designer can set up the appliances according to the changes occurring to the dwellers and their environment. In the SHS, the combination of Virtual Reality and Semantic Web generates an integrated and aggregated view on relevant knowledge related to the domestic environment. This coupling allows physical and virtual objects to coexist and interact in real time, thus enhancing appliances' configuration operations also in real scenarios. Moreover, Semantic Web technologies can provide a formal method to represent and model the digital counterpart of the real physical world corresponding to the domestic environment.

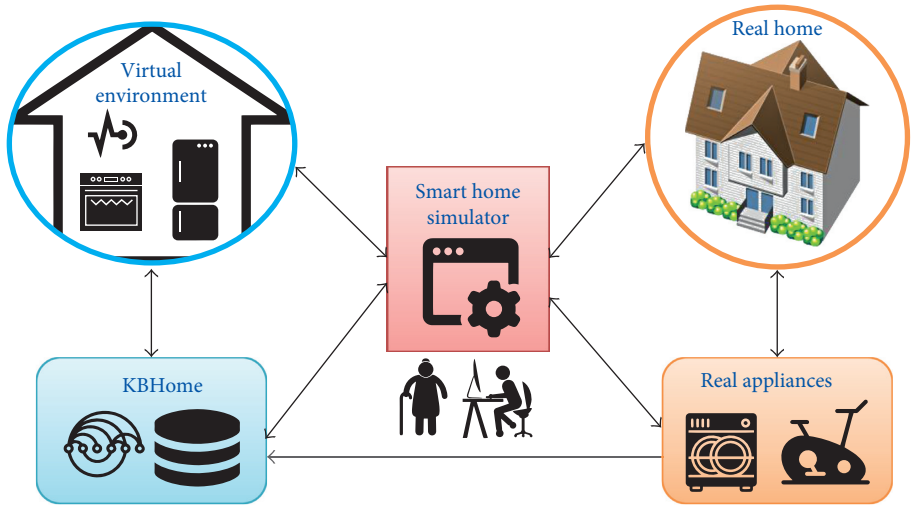


Figure 2: The role of the Smart Home Simulator.

3.5. The System's Architecture. The SHS is an application relying on an integrated service-oriented platform described in [45]. The main purposes of this platform are to manage the knowledge about the home and user's domains, while allowing the various appliances to exchange data among them. This framework is called Virtual Home Framework and is based on four main pillars:

(i) The first is the semantic layer, named "KnowledgeBase Home" (KBHome), which is the set of ontologies describing the relevant knowledge of the abovementioned domains (the KBHome, further described in Section 3).

(ii) The second pillar consists in the virtual representation of the domestic environment: this feature allows to virtually represent appliances and sensors (as described in Sections 4 and 5).

(iii) Integration Services guarantee data integration synchronization between physical and virtual components of the system. This approach is delved in [46], and in this context, applications can easily interoperate while the data are integrated, shared, dispatched, and aggregated through mechanism transparent to their clients. Therefore, the Integration Services are able to promote the semantic integration among the data provided by the various domestic appliances and sensors (both real and virtual) and to contribute to enhance their near real-time synchronization:

(1) They enable the acquisition of information from any device (appliance or sensor and real or virtual).
(2) They allow to store, interpret, and manage the information received.
(3) They share and dispatch information when a device asks for them or when a needed information becomes available.

(iv) The Real Home, with its real appliances, represents the deployment of the solutions identified, thanks to the cooperation among the three pillars of the framework.

The SHS acts like a user-friendly, semantic, and virtual reality-based interface to allow designers to configure and test (with either virtual or real appliances) AAL solutions customized on specific users (Figure 2).

4. The Home Knowledge Base

The knowledge base containing the information of the user and his/her health and physical status and the living environments composing the house and the appliances deployed or deployable in it are modelled in a set of ontologies named "KBHome." The ontologies are modelled following the NeOn Methodology [47], an ontology engineering methodology allowing to focus on the reuse of already existing resources—both ontological or not. The knowledge base has been developed using the software Protégé [48], while Resource Description Framework (RDF) [49] and Web Ontology Language (OWL) [50], with the use of Semantic Web Rule Language (SWRL) [51], were the selected implementation languages; Pellet [52] was the reasoner used to perform reasoning activities on the ontologies. KBHome is composed of four main ontologies, each addressing a specific domain.

4.1. The User's Health Condition and Cardiorespiratory Fitness Ontology. This domain ontology describes concepts regarding the domain of the user, starting with his registry records (date and place of birth, current address, gender, and phone numbers). The most relevant feature of this ontology is the description of the user's health condition using the codes and qualifiers included in the ICF. To this purpose, the ontology of ICF—publicly available on the BioPortal—was partially reengineered: the specific ICF codes, originally modelled as individuals, were converted into datatype properties, in order to make possible to model several health conditions using the same ICF code. Each user is linked to his/her health condition (modelled as an individual), which

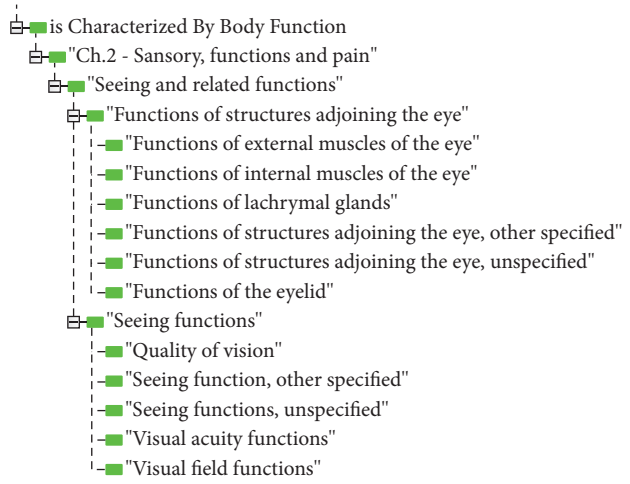


Figure 3: An excerpt of ICF "Body Functions" in the KBHome.

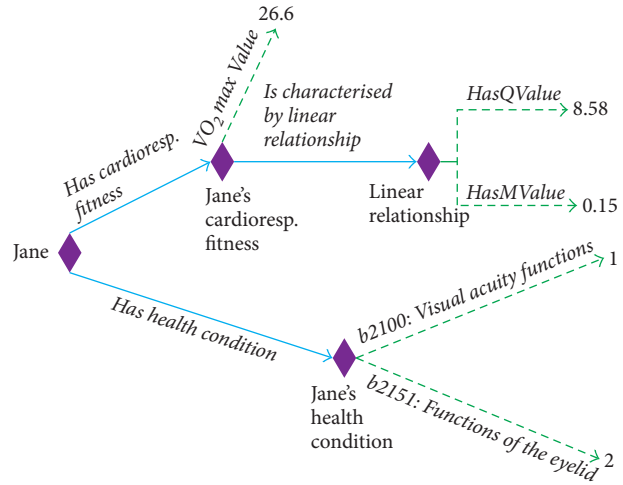


Figure 4: An example of user characterization in the KBHome.

is described via the datatype properties composing the ICF (Figure 3).

Each health condition can then be classified as "Vision Impaired Health Condition," "Motor Impaired Health Condition," "Hearing Impaired Health Condition," and "Cognitive Impaired Health Condition." For each of these classes, there is a further subclassification regarding the grade of impairment associated, according to the qualifier assigned to each datatype property; for instance, the class "Vision Impaired Health Condition" can be deepened into "Mild Visual Impairment," "Moderate Visual Impairment," "Severe Visual Impairment," and "Complete Visual Impairment." According to the typology of impairment associated with his/her health condition, a user can then be inferred to belong to a specific class of users ("Vision Impaired User," "Motor Impaired User," "Cognitive Impaired User," or "Hearing Impaired User").

Following the same ontology design pattern, a user is linked to his/her CRF (indicated as $\dot{V}O_2max$), which is assessed during a test under the supervision of clinical personnel. According to the value detected and stored in this ontology, the user's CRF can be classified, thanks to reasoning processes into subclasses ("Seriously limited CRF," "Reduced CRF," "Adequate CRF," "Good CRF," "Excellent CRF," and "More than excellent CRF," according to the percentile [53] the user's value fits).

User's $\dot{V}O_2max$ value can be used to assess the customized workload of an exercise (in this case performed on a cycle-ergometer), as described in the following equation [54]:

$$\text{WL (cycle} - \text{ergometer)} = \frac{\left(\left(\dot{V}O_2\text{max} \cdot 0.65\right) - q\right)}{m}. \quad (1)$$

The angular coefficient m and the intercept q are user-dependent and are calculated from the interpolation of the data obtained during the test execution. A training intensity of 65% is chosen in agreement with ACMS, which identifies it as the minimum intensity capable of producing an increase in the user's CRF [55]. Equation (1) can be easily translated into SWRL, thus determining the exercise load to be applied to the cycle-ergometer:

```
Device(?cycle), User(Jane),TrainingIntensity
(?trInt), hasTrIntValue (?trInt, ?TrInt-
Value), hasCardiorespFitness(Jane, ?cf),
isCharacterizedByLinearRelationship(?cf,
?lr), hasMValue(?lr, ?mvalue), hasQValue(?lr,
?qvalue), hasVO2maxValue(?cf, ?VO2max), mul-
tiply(?step1, ?VO2max, ?TrIntValue), subtract
(?step2, ?step1, ?qvalue), divide(?WLcycle,
?step2, ?mvalue), → setsWLonCycleErg(?cycle,
?WLcycle)
```

In this way, it is possible to represent relevant knowledge about the user and some features of his/her physiological status in a simple way, as summarized in Figure 4.

4.2. Appliances and Domestic Environment Ontology. This ontology aims at providing a description of the appliances and linking them to the room where they are deployed. Moreover, it allows to provide a list of the measurements each appliance can perform. The ontology is composed of three main modules: the first provides a simple representation of a generic domestic environment, modelled with classes and individuals. Each individual represents a room of the user's house and is described with datatype properties illustrating the dimensions of each room.

The second module is the appliances module, which collects a list of appliances (both white and brown goods) and sensors and provides a description for each of them; this description is achieved taking advantage of the HicMO "grammar" [56], a set of XML properties able to describe the features of any appliance. According to the guidelines provided by NeOn methodology, HicMo properties were converted into datatype and object properties to create a semantic model of the XML descriptors. In this way, it is possible to provide a sort of "ID card" for each appliance deployed or deployable in the house. The description of a Smart Object is integrated with a submodule describing the

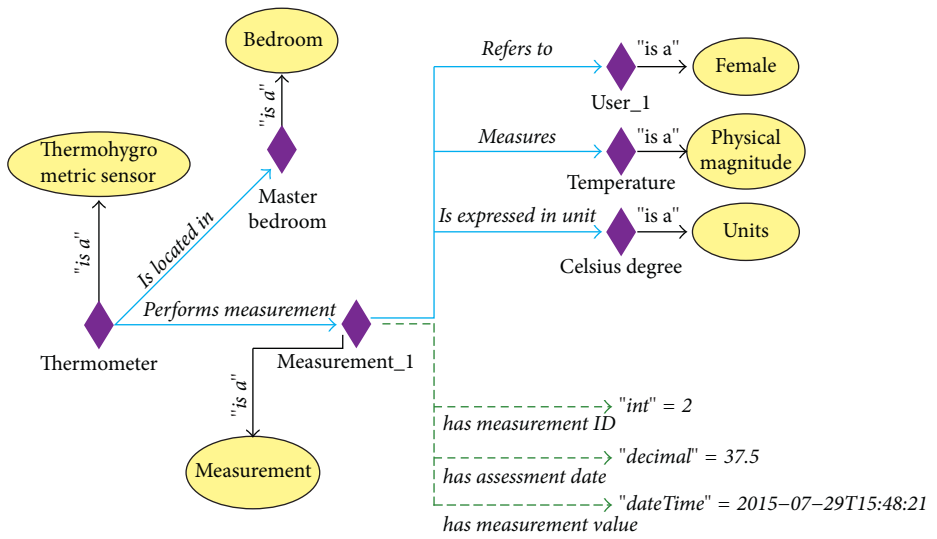


Figure 5: An example of appliance modelled in the KBHome.

list of programs available for each appliance; each program is described as an individual and is associated with one or more appliances.

Since each appliance can produce one or more measurements, the third module allows to provide a description of these measurements. The latter can be classified into "Environmental measurements," "Vital sign," and "Appliance measurements"; for instance, a digital thermometer located in the living room can detect the temperature of the room, while a thermometer located on the user measures his/her internal temperature. Measurements describing the user's physiological status exploit a partially reengineered version of the Vital Sign Ontology [57], while measurements regarding appliances and domestic environments exploit the Units Ontology [58] to provide a sound description. Figure 5 depicts an example of a sensor performing a measurement of the user's temperature.

4.3. Comfort Metrics and Domestic Environment Model. This ontology describes comfort dimensions inside the domestic environment, taking into account measurements modelled in classes such as "CO_2Concentration," "DomesticHumidityRate," "DomesticLuminance," and "DomesticTemperature." In particular, the class "DomesticTemperature" is divided into "WinterDomesticTemperature" and "SummerDomesticTemperature." Each of the classes is split into "Acceptable" and its complement "NotAcceptable." Each individual "Environmental measurement" is then classified according to these classes. For instance, a measurement detecting a CO_2 concentration equal to 1147 ppm is inferred to be a "NotAcceptableCO_2Concentration," since the "AcceptableCO_2Concentration" requires environmental measurements with a value less than or equal to 1000 ppm.

The classes for the description of comfort dimensions convert the limits described into several standards (such as ASHRAE [59] for the thermal and humidity rate comfort, UNI-2004 [60] for the luminance recommendations, and UNI-2008 [61] for the air quality in domestic environments).

4.4. Orchestration of Services in the House. This ontology describes the events triggered by one or more of the conditions occurring in the environment or to the user. With the use of SWRL rules, it is possible to describe the conditions under which a specific action is activated. For instance, to set the proper air-conditioning system's program during summer, the following conditions must hold:

```
EnvironmentalMeasurement (?m), NotAcceptableSummerDomesticTemperature (?m), hasMeasurementValue (?m, ?value), greaterThanOrEqual (?value, 27), EnvironmentalMeasurement(?n), NotAcceptableDomesticHumidityRate (?n), hasMeasurementValue (?n, ?value2), greaterThanOrEqual (?value2, 60), Appliance (air_conditioning_system), AirConditioningSystemProgram (SummerBreeze) -> hasProgram (air_conditioning_system, SummerBreeze)
```

Several situations, similar to the one described above, were described in this ontology, providing the conditions under which the appliances are able to respond with the proper services. More complex situation involves customized services to be deployed for particular categories of impaired users.

4.5. Configuration of Living Environments. This ontology allows to classify the appliances (described as noted in Section 3.2) into other classes; according to their characteristics and the programs they have, appliances can provide useful services to the users, to help them cope with their impairments while performing activities of daily living. Therefore, the appliances of the house are classified also according to their suitability for specific categories of impaired users, using the classes "Cognitive Impairment Appliance," "Hearing Impairment Appliance," "Motor Impairment Appliance," and "Visual Impairment Appliance."

Taking advantage of these features, it is possible to model a configuration project. Each project is modelled as an

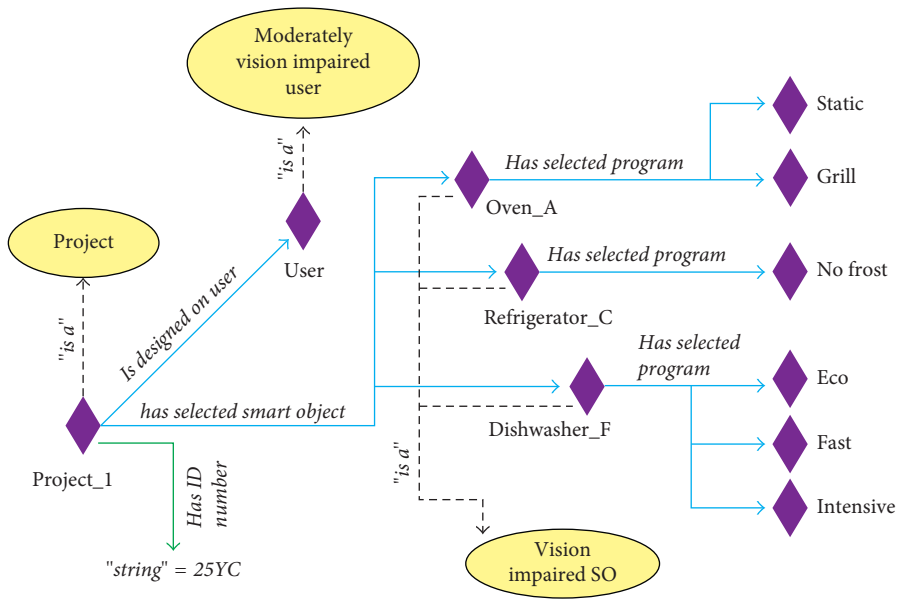


Figure 6: A project modelled in the KBHome for a vision-impaired user.

individual with an ID number and allows to select appliances together with their programs. Furthermore, it is possible to specify the user for which the project is designed (Figure 6). This ontology also allows to automatically infer if an appliance is suitable for a particular category of impaired users according to the programs available for the application. In this way, a designer can choose from a list of appliances inferred to be adequate to cope with a user's impairment, and the outputs coming from the reasoning process of this ontology are exploitable as a decision support system [62] for the design phase of the living environment.

5. Customization of the Smart Home Services

In order to design a Smart Home able to provide its dwellers with the right instruments and services that enhance their well-being and autonomy in ADL, the customization of the offered solution is necessary. This step passes through the collection of the subjects' needs—through the evaluation of their health status—and the design of a customized solution able to help the users in coping with their impairments or to improve in some way their quality of life.

5.1. User Characterization. For the assessment of the health condition of the home dweller(s), the intervention of clinicians is required to fill in the ICF-based module. The clinical personnel have to assess all the user's impairments through standard tests and clinical scales specifically dedicated to investigate a precise domain (vision, hearing [63], motor or cognitive deficits) or the general health status in elderlies [64, 65].

The evaluation of the CRF occurs in a second phase and can be performed in different ways. The gold standard methodology consists in the measurement of oxygen and carbon dioxide through spirometry during a physical exercise of incremental effort. This leads to the determination of the maximum oxygen uptakedefined as $\dot{V}O_2$— which represents the ability of an individual to take up and use the inspired oxygen [22]. Alternative techniques foresee the indirect estimation of the CRF starting from the measurement of the subject's heart rate (HR). These techniques are based on the assumption of a linear relationship between the user's HR and the workload; moreover, they assume that all individuals of the same age have indeed the same maximal cardiac frequency. Though these hypotheses are usually true for healthy subjects, these assumptions often do not stand in the case of elderlies or people with disability [23]. Thus, the first methodology—based on direct measurements of expired gases—should always be preferred. Another important aspect to be taken into account is the safety of the tested individual during the CRF evaluation test. The American Thoracic Society identified the cycle ergometer as the equipment to be preferred with respect to the treadmill in case of patients, because it does not introduce risk of falls and allows the concurrent measurement of gases and work rate, in exchange of a reduced capability of reaching high levels of $\dot{V}O_2$, which should not be an issue in case of frail individuals.

Having defined the equipment and other supplementary tools—a blood pulse oximeter, an electrocardiograph (ECG), and a blood pressure (BP) monitor at least—needed to ensure the safety of the subject during the test, the CRF of the Smart Home dweller can be identified following a multistage exercise protocol [22]. The chosen protocol foresees a preliminary warm-up phase with low and constant workload followed by an increase of 20 W in the workload after a steady state of 3 minutes. All the conditions triggering the increase in workload are reported in details in Table 1.

During the test, expert clinical personnel must continuously monitor the conditions of the individual performing the cardiopulmonary exercise: this is necessary since the final aim is to identify the subject's maximal effort reaching

Table 1: Parameters to conduct the maximal cycling exercise test.

HR warm-up (bpm)	Starting WL (W)	ΔWL (W)	Max WL (W)
≤80	100	+20	160
80 < HR ≤ 90	80	+20	140
90 < HR ≤ 100	60	+20	100
≥100	40	+20	80

Table 2: Conditions for interrupting the maximal cycling exercise test. These indications were adapted from [22] introducing more strict criteria to prevent the onset of critical conditions in frail elderlies.

Conditions for test interruption
Chest pain suggesting ischemia
ECG changes suggesting ischemia or ectopy
Heart rate exceeding 85% of maximum estimated HR ($HR > 0.85\ (220 - age - HR^{rest}) + HR^{rest}$)
Fall in blood pressure ($BP_s < BP_s^{rest}$, $BP_d < 80$ mmHg)
Hypertension ($BP_s > \min\{200$ mmHg, $BP_s^{rest} + 10$ mmHg$\}$ or $BP_d > 110$ mmHg)
Severe desaturation ($SpO_2 < 80\%$)
Sudden pallor
Dizziness, faintness, or confusion
Loss of limb coordination
Dyspnoea

HR = heart rate, BP_s = systolic blood pressure, BP_d = diastolic blood pressure, SpO_2 = arterial oxygen saturation.

his/her physiological limits. Therefore, if any of the conditions listed in Table 2 occurs, the exercise must be immediately terminated. Moreover, in case of exercise interruption, the subject must be assisted by a physician until normal values and a stable condition are recovered.

Having completed these two phases of evaluation, each subject is provided with a complete record of his/her health status—according to the worldwide standard of ICF—accurate information about his/her CRF, and thus his/her physical capabilities. From the latter, the customization of the daily physical exercise for each individual could be addressed (as described in Section 6.3). The cycle ergometer is maintained as the training equipment also in the Smart Home environment configuration because of its higher safety with respect to the treadmill, which is the only other fitness equipment able to provide a direct control on the workload. Of course, the use of a cycle ergometer is compatible with individuals who do not suffer from severe motor limitations or have impairments in the postural control. These severe situations are identified during the first clinical assessment, formalized through the ICF completion and should be addressed with customized solutions, as for instance, an arm ergometer.

5.2. Configuration and Test of the User of the Smart Home. Starting from the pieces of information gathered during the user's evaluation phase, the Smart Home designer, being able to interpret these data, can set up a mixed reality environment able to respond to the user's needs with the most appropriate solutions. In this context, using either mixed or virtual reality leads to several advantages; the first, as already mentioned, is the possibility to test the communication among the different appliances in real time. The second consists in the possibility for the final user to directly experience the VEs, with the double aim of becoming familiar with his/her newly configured Smart Home and of giving suggestions to the designer who can improve the final home design. Finally, the employment of virtual reality, coupled with the semantic models described in Section 3, allows the Smart Home designer/the final user to test the functioning of a sensor or an appliance without the need of owing it in real world, thus saving time and costs.

5.2.1. Customized Configuration of the Smart Home. In this phase, the designer receives the blueprint of the final users' current house; he/she is thus able to model it using an authoring tool, reconstructing a digital model of the dwelling rooms. This model can then be imported inside the Smart Home Simulator, the PC-based application developed in the D4All project using Unity 3D. Within this application, the designer can add devices and appliances to the house digital model, choosing them among the ones modelled in the semantic "catalogue" described in Section 3.2. The catalogue is updated in real time inside the VE using SPARQL to query the semantic repository, in which the user's needs and peculiarities were stored.

The designer has also the possibility to integrate real devices or sensors in the digital representation of the Smart Home, thanks to the architecture described in Section 3.5, which permits the data exchange between the real and the virtual world. When the design of the customized environments is completed and the communication among the devices, the sensors, and the human users is validated, the designer can save the project and store it for further modifications.

5.2.2. Testing the Designed Solution. At the end of the design phase, the Smart Home Simulator offers the possibility for the final user to test the solution specifically developed according to his/her characteristics and needs. The solution can be deployed on different platforms, depending either on the characteristics of the target user or on the type (virtual or mixed reality) of environment. Of course, immersive (e.g., head-mounted displays and CAVE) or semi-immersive (e.g., semicylindrical projected screens) experiences constitute the most promising means to validate different scenarios and to learn how the Smart Home services work, because of the higher sense of presence they elicit and the more natural interaction they often provide [66]. However, when choosing the VR technologies, particular attention must be paid on the target user: for frail elderlies or severe cognitive impaired subjects, the risks of adverse events and sickness while using head-mounted displays are not deniable [65]. Therefore, other solutions should be preferred, even in exchange of a reduced sense of presence. The use of non-immersive environments should also be preferred in case of

real devices or appliances included in the final solution, because—if not properly reconstructed in the virtual world—the interaction may become very complicated or even dangerous (i.e., the user can trip).

After the validation in the mixed reality environment, the designer is able to determine if the implemented solution is able to help the user in coping with his/her limitations during activities of daily living, having observed his/her behaviour during the simulation. Moreover, he/she can retrieve direct information from the final user, interviewing him/her and asking about the changes they would make in their future house. These modifications can be then implemented in the configurator and retested till the reaching of the optimal solution that the final user will implement at his/her own place.

6. Validation with Real Use Cases

The SHS of the adopted framework were validated through two real use cases. The first foresees the configuration of a kitchen for a visually impaired user, while the second addresses the problem of active ageing and foresees an elderly user performing domestic physical activity.

6.1. Configuration of a Real Environment: The Kitchen. This use case consists of a kitchen designer who has to design a kitchen using the SHS selecting the most suitable appliances for the final user: a person afflicted by a moderate visual impairment (specified with ICF codes: b21022.2—"moderate impairment in the contrast sensitivity") and hyposmia (b255.2—"moderate impairment in the smell function"). The designer acquires the user's kitchen blueprints and elaborates a virtual model of the kitchen. In this model, he/she places the most suitable appliances and sensors to cope with the user's impairment(s) choosing them from a list, provided by the KBHome (as seen in Section 3.5).

For this user, the KBHome retrieves

(i) two induction cooktops: one able to produce a high contrast on its surface (with a black panel) with textured button surfaces; the other providing a black surface with high-contrast controls and digital and backlit display;

(ii) two models of convection ovens with digital and backlit display and a set of control lights;

(iii) four models of dishwashers with high contrast and backlit digital display;

(iv) two models of refrigerators with digital and backlit interface, illustrating the current inside temperature and internal light.

According to the user's preferences, the designer chooses the appliances and sets them into the virtual model of the kitchen, as illustrated in Figures 7–10.

6.2. Validation of the Application. The validation of the configuration part is performed following the formative evaluation methodology, as described in [67, 68]. This type of validation foresees an observation and empirical evaluation of representative users' interaction with the VEs in a task-based scenario: in this case, the design of the kitchen described in Section 6.1. The aims of this type of evaluation are identifying usability problems and assessing the user's learning curve and his/her task performance, both with qualitative and quantitative results. To do this, five designers will be enrolled in the study; they will be presented the scenario and the Smart Home Simulator application during

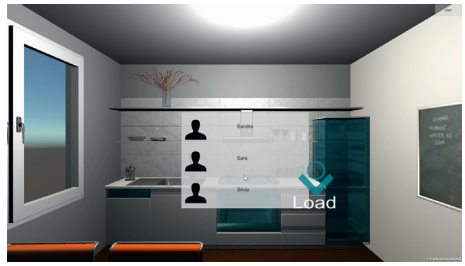


Figure 7: A snapshot of the Smart Home Simulator application: the designer selects the user for who he/she configures the kitchen.

Figure 8: A snapshot of the configurable appliances (for a selected user) in the kitchen.

Figure 9: The designer places the dishwasher, choosing from the appliance models retrieved from the KBHome.

Figure 10: The designer places the suitable sensors in the kitchen.

a first training phase of about 15 minutes. To ease and speed up their work, the digital model of the empty kitchen will also be provided as the starting point. During the configuration phase, tasks timing, errors, potential software bugs, and users' comments will be registered. At the end of the trial, the System Usability Score (SUS) questionnaire [69] will be administered to each participant. Free comments will also be collected and taken into account to improve the software according to the final users' (the designers) needs.

6.3. Configuration of Living Environment for Physical Activities. The first use case depicts a common situation for elderly people, where a 72-year-old frail woman has to perform daily the physical exercise on a cycle-ergometer placed in her bedroom. Her health condition is described by the following ICF codes: b7353.2 ("moderate impairment in the tone of muscle of lower half of the body"), b4550.2 ("moderate impairment in the general physical endurance"), and s75002.181 ("mild impairment to the muscle of the right tight (nature of the impairment not specified)), while her CRF was assessed at 18.3. Therefore, she has been prescribed an adequate physical activity, to be performed monitoring her results and conditions. The data deriving from her exercise sessions must be made available to caregivers or clinical personnel, in order to periodically assess enhancements or deteriorations of her abilities and conditions. The use case foresees that the user's smart home is equipped with the cycle ergometer, a heart rate monitor, a breath rate monitor, and a blood pressure monitor. The user receives information about the performance by means of a tablet. To accomplish her daily physical activity, she is requested to enter the bedroom, where a presence sensor detects the user's presence. She is then required to prepare the bedroom in order to perform the exercise, by ventilating the room and waiting until it reaches the proper temperature (which is set at 20°C). Once the air quality and the temperature reach the proper values, the user can wear the sensors and begin the exercise, whose workload is set basing on her specific health condition. While performing the activity, her physiological status is monitored. At the end of the physical exercise, the data regarding the exercise session and the physiological measurements detected are stored and remain available for the caregivers. All the instructions are given to the user via a virtual tablet, through which she can also check her performance in real time and potentially receive different types of alerts based on the sensors' measurements.

The designer, using the SHS features, equips the environment with real sensors (heart rate monitor, breath rate monitor, and blood pressure meter) and provides the virtual representation of the living environment with an actuator to enable the automatic opening and closing of the window. The virtual environment also includes a presence sensor, able to detect the user's presence in the room, an environmental thermometer, and an air-quality sensor, able to measure the CO_2 concentration in the room. Data regarding the user's health condition and CRF are stored in the KBHome.

By navigating the virtual scene, the user taps the virtual tablet on the exercise icon and receives the instructions about how to set the environment for her activity. At first, she is asked to enter the bedroom, where the cycle ergometer is located. The presence sensor is able to detect her presence and triggers the air quality sensor installed in the room. It measures the CO_2 concentration in the room and stores the acquired data into the KBHome, where this piece of information is processed. Hence, having registered an unacceptable CO_2 concentration in the room, the user receives an alert on the tablet warning her that, due to the current CO_2 concentration, she is not allowed to perform the exercise: the user receives the suggestion to open the window. Then, she opens the window by tapping the proper option on the virtual tablet. When the air quality and the temperature reach a suitable level to perform the exercise, the user receives a notification on the tablet. Finally, she can begin the physical exercise, whose duration and level of difficulty are automatically selected by the KBHome based on her health condition and CRF assessment (as shown in Sections 4.1 and 5.1). In this case, the workload is calculated according to the equation presented in Section 4.1, and the duration of the exercise is set at 20 minutes (or user request). The user receives instructions on how and where to wear the sensors. While performing the exercise, the user's blood pressure, breath rate, and pulse rate are monitored by these wearable devices, and, if any physiological anomaly arises, the user is warned to immediately modify or stop the exercise via a tablet alert. At the end of the exercise, physical and physiological data regarding the performance (blood pressure, pulse rate, and duration of the physical activity) are stored in the Semantic Repository, where they are at the caregivers' disposal.

6.4. Development of the Validation on Elderly Subjects for the Configured Environment for Physical Activities. The validation on the elderlies follows the same methodology described for the kitchen design scenario. A small sample ($n = 5$) of target users will be enrolled in a first-pilot trial dedicated to the assessment of potential software or methodological issues in the developed scenario.

Enrolled subjects will have to fulfil the following inclusion criteria: age ≥ 65 years old, they should have a mild-to-moderate impairment in their general physical endurance, and they should be judged by a clinician as subjects who would benefit from a light daily physical exercise. Exclusion criteria are the presence of severe cognitive deficits or vision problems and the inability to express the informed consent. Each of the subjects, once enrolled, should undergo a CRF assessment in the clinical setting, as described in Section 5.1. After the completion of the test, his/her CRF value will be used to assess the target workload to be set at the beginning of the exercise.

The setup used for the validation of the scenario dedicated to elderlies tries to mimic as much as possible the situation described in Section 6.3. Thus, an entire room is completely dedicated to the recreation of the described scenario; it will be provided with a cycle ergometer (real), a pulse oximeter (real), a blood pressure meter (real), a tablet (real), a presence sensor (real), an automated window

Table 3: Conditions for interrupting the domestic exercise of the user cycling exercise test. These indications were adapted from [22] introducing more strict criteria.

Conditions for exercise interruption
Heart rate exceeding 85% of maximum estimated HR ($HR > 0.85\ (220 - age - HR^{rest}) + HR^{rest}$)
Fall in blood pressure ($BP_s < BP_s^{rest}$, $BP_d < 80$ mmHg)
Hypertension ($BP_s > \min\{200\text{ mmHg}, BP_s^{rest} + 10\text{ mmHg}\}$ or $BP_d > 110$ mmHg)
"Very strong" perceived effort (Borg Category-Ratio Scale [70] (CR10) ≥ 7)
"Severe" pain (VAS Pain Scale [71] ≥ 7)

HR = heart rate, BP_s = systolic blood pressure, BP_d = diastolic blood pressure, SpO_2 = arterial oxygen saturation.

(virtual), an air quality sensor (virtual), and a thermometer (virtual). All the virtual objects will be presented to the target user using a wall projector.

Expert personnel will instruct each subject about the aim and the functionalities of the system, providing an overview of each component. As soon as they are confident with the setup, the subjects will be left free—under constant supervision by clinical and technical personnel—to interact with the system following the instructions given through the tablet. Each subject will perform the 20 minutes of cycling, after having opened and closed the (virtual) window, with the workload set according to his/her CRF. In order to guarantee, at each instant, the safety of the training, the monitored parameters must not exceed the values reported in Table 3.

The validation of this scenario will pass through the collection of objective quantitative data (errors, task timing, need of suggestions, and general performance) and the interview of the subjects using an ad hoc developed questionnaire. Due to the characteristics of the population, in fact, semistructured interviews were chosen as the preferred methodology to gather qualitative information assessing the acceptability and usability of the designed system [72].

7. Conclusion and Further Works

This paper presents the Smart Home Simulator, an AAL application that takes advantage of both Virtual Reality and Semantic Web technologies to tackle the configuration of domestic environments. The dwellers' health conditions, periodically assessed by clinical personnel, are modelled into a semantic knowledge base (KBHome) that allows to automatically infer a set of appliances able to help the dwellers in performing several daily life activities autonomously. The ontologies also provide a formal description of the users' health conditions and their cardiorespiratory fitness, environmental comfort metrics, and appliances and their behaviours enabling the possibility to provide tailored services to the dwellers. In fact, the results of reasoning process allow, from the one hand, to identify a set of appliances that can support the dwellers in daily life activities, thus helping them to cope with their impairments and to live in an autonomous way; on the other hand, KBHome can provide the dwellers with parameters to set up a daily physical activity. One of the most relevant aspects of the Smart Home Simulator relies on the exploitation of the International Classification of Functioning, Disability and Health, an international standard acting as a common language in health-related fields among clinical personnel and nonclinical professionals.

The results coming from the reasoning process are fed to a Virtual Reality-based application, able to provide a reproduction of a real domestic environment in which smart home designers or architects can select and deploy the appliances to study their behaviours and how they can cope with users' impairments. This virtual configuration process has the advantages of being completely organized around user's real needs and limits (modelled into the ontology with ICF) and to considerably reduce the costs and time to set up a smart home by testing its appliances (and their behaviours) in a virtual environment.

Two use cases were deployed to validate the Smart Home Simulator functionalities and usability and to test the possibility to configure a specific environment for performing physical exercise according to particular criteria.

In the next months, the validation phase described in Section 6 will take place involving different medical and industrial partners. The validation process includes experimental campaigns with patients (characterized by mild and moderate impairments), who will be asked to dive into the Virtual Environments and to perform the tasks described in the use case provided in Section 6.3. In this way, it will be possible to validate the methodological approach and the chosen technologies.

Another goal of the validation phase is to optimize the framework in order to ensure comfort levels to elderly people inside their homes and to extend the approach here described to the measurement and the assessment of domestic comfort indicators.

Further works foresee the development of other scenarios involving different kinds of users, characterized by diverse impairments and different domestic environments. The KBHome ontologies will be enriched to comprise a larger number of comfort metrics and to model rules to provide tailored comfort metrics to specific categories of dwellers. These developments are expected to be eased by the scalar architecture already implemented and will allow to provide suitable answers to various users' needs, as foreseen in the "D4All" project.

Another improvement of the presented approach will also be addressed in the future. An attempt to partially automatize the process of information retrieval from the semantic models and to subsequently exploit those data with Machine Learning (ML) techniques will be performed. The use of ML techniques should be integrated into the semantic approach, allowing the modification of the knowledge base. In this case, some applications should be developed to ensure the proper modelling of the data acquired via ML techniques; otherwise, there exists the possibility to corrupt the original ontologies. Furthermore, using the ML approach, the supervision of an expert (e.g., a designer or a clinician), that today is strongly necessary, may be limited

only to crucial tasks (e.g., decision-making on the users' health). Indeed, the complete automation of the process still represents an open challenge, because of the sensitive domains that could be affected by wrong decisions taken by the ML algorithms. When dealing with (frail) human users and their health, in fact, no mistakes could be tolerated, and, since nowadays there is no guarantee that automatic learning always works perfectly, the supervision of a human expert is required for both ethics and legal reasons.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to acknowledge the Progetto MIUR no. CTN01_00128_297089 "Design for All—Sw integration and advanced Human Machine Interfaces in design for Ambient Assisted Living."

References

[1] Q. Ni, I. Pau de la Cruz, and A. B. García Hernando, "A foundational ontology-based model for human activity representation in smart homes," *Journal of Ambient Intelligence and Smart Environments*, vol. 8, no. 1, pp. 47–61, 2016.

[2] L. Razmerita, A. Angehrn, and A. Maedche, "Ontology-based user modeling for knowledge management systems," in *Proceedings of the International Conference on User modeling (UM 2003)*, p. 148, Johnstown, PA, USA, June 2003.

[3] J. Mocholí, P. Sala, C. Fernández-Llatas, and J. Naranjo, "Ontology for modeling interaction in ambient assisted living environments," in *Proceedings of the XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*, pp. 655–658, Chalkidiki, Greece, 2010.

[4] M. Golemati, A. Katifori, C. Vassilakis, G. Lepouras, and C. Halatsis, "Creating an ontology for the user profile: Method and applications," in *Proceedings of the first International Conference on Research Challenges in Information Science (RCIS 2007)*, pp. 407–412, Ouarzazate, Morocco, April 2007.

[5] K.-L. Skillen, L. Chen, C. Nugent, M. Donnelly, W. Burns, and I. Solheim, "Ontological user profile modeling for context-aware application personalization," in *Proceedings of the International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2012)*, pp. 261–268, Vitoria-Gasteiz, Spain, 2012.

[6] M. Sutterer, O. Droegehorn, and K. David, "Upos: user profile ontology with situation-dependent preferences support," in *Proceedings of the First International Conference on Advances in Computer-Human Interaction (ACHI 2008)*, pp. 230–235, Sainte Luce, Martinique, 2008.

[7] R. Welge, B.-H. Busch, K. Kabitzsch et al., "Representation of integration profiles using an ontology," in *Proceedings of the Ambient Assisted Living*, pp. 195–212, Springer, Berlin, Heidelberg, 2014.

[8] G. Guerrero-Contreras, J. L. Navarro-Galindo, J. Samos, and J. L. Garrido, "A collaborative semantic annotation system in health: towards a SOA design for knowledge sharing in ambient intelligence," *Mobile Information Systems*, vol. 2017, Article ID 4759572, 10 pages, 2017.

[9] J. L. Navarro-Galindo and J. Samos, "The FLERSA tool: adding semantics to a web content management system," *International Journal of Web Information Systems*, vol. 8, no. 1, pp. 73–126, 2012.

[10] A. Grguric, D. Huljenic, and M. Mosmondor, "AAL ontology: from design to validation," in *Proceedings of the IEEE International Conference on Communication Workshop (ICCW 2015)*, pp. 234–239, London, UK, 2015.

[11] E. Meshkova, J. Riihijarvi, P. Mahonen, and C. Kavadias, "Modeling the home environment using ontology with applications in software configuration management," in *Proceedings of the International Conference on Telecommunications (ICT 2008)*, pp. 1–6, St. Petersburg, Russia, 2008.

[12] T. Van Nguyen, J. G. Kim, and D. Choi, "ISS: the interactive smart home simulator," in *Proceedings of the 11th International Conference on Advanced Communication Technology (ICACT 2009)*, vol. 3, pp. 1828–1833, Phoenix Park, South Korea, 2009.

[13] J. Park, M. Moon, S. Hwang, and K. Yeom, "CASS: a context-aware simulation system for smart home," in *Proceedings of the 5th ACIS International Conference on Software Engineering Research, Management & Applications (SERA 2007)*, pp. 461–467, Busan, South Korea, August 2007.

[14] P. Sernani, D. Calvaresi, P. Calvaresi, M. Pierdicca, E. Morbidelli, and A. F. Dragoni, "Testing Intelligent Solutions for the Ambient Assisted Living in a Simulator," in *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA'16)*, p. 71, Corfu, Island, Greece, July 2016.

[15] A. Helal, K. Cho, W. Lee, Y. Sung, J. Lee, and E. Kim, "3D modeling and simulation of human activities in smart spaces," in *Proceedings of the 9th International Conference on Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC 2012)*, pp. 112–119, Fukuoka, Japan, September 2012.

[16] J.-W. van't Klooster, B.-J. van Beijnum, A. Eliens, and H. Hermens, "Interactive scenario visualization for user-based service development," in *Proceedings of the International Conference on Collaboration Technologies and Systems (CTS 2012)*, pp. 498–503, Denver, CO, USA, May 2012.

[17] C. Fernández-Llatas, J. B. Mocholí, P. Sala et al., "Ambient assisted living spaces validation by services and devices simulation," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, 2011*, pp. 1785–1788, Boston, MA, USA, August 2011.

[18] G. Edlinger, C. Holzner, C. Guger, C. Groenegress, and M. Slater, "Brain-computer interfaces for goal orientated control of a virtual smart home environment," in *Proceedings of the 4th International IEEE/EMBS Conference on Neural Engineering (NER'09)*, pp. 463–465, Antalya, Turkey, 2009.

[19] C. Guger, C. Holzner, C. Grönegress, G. Edlinger, and M. Slater, "Control of a smart home with a brain-computer interface," 2008.

[20] R. L. Mace, G. J. Hardie, and P. Jaine, "Place accessible environments: toward universal design," in *Design Interventions: Toward A More Humane Architecture*, Van Nostrand Reinhold, New York, NY, USA, 1991.

[21] R. Costa, D. Carneiro, P. Novais et al., "Ambient assisted living," in *Proceedings of the 3rd Symposium of Ubiquitous Computing and Ambient Intelligence 2008*, pp. 86–94, San Sebastián, Spain, 2009.

[22] M. Pieper, M. Antona, and U. Cortés, "Ambient assisted living," Ercim News, vol. 87, pp. 18–19, 2011.

[23] H. Sun, V. De Florio, N. Gui, and C. Blondia, "Promises and challenges of ambient assisted living systems," in *Proceedings of*

the Sixth International Conference on Information Technology: New Generations, 2009 (ITNG'09), pp. 1201–1207, Las Vegas, NV, USA, 2009.

[24] D4All, Sw Integration and Advanced Human Machine Interfaces in Design for Ambient Assisted Living (project ID: CTN01_00128_297089), 2017, http://www.d4all.eu.

[25] R. Harper, *Inside the Smart Home*, Springer Science & Business Media, London, UK, 2003.

[26] S. Katz, "Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living," *Journal of the American Geriatrics Society*, vol. 31, no. 12, pp. 721–727, 1983.

[27] D. E. R. Warburton, C. W. Nicol, and S. S. D. Bredin, "Health benefits of physical activity: the evidence," *Canadian Medical Association Journal*, vol. 174, no. 6, pp. 801–809, 2006.

[28] A. Clegg, J. Young, S. Iliffe, M. O. Rikkert, and K. Rockwood, "Frailty in elderly people," *The Lancet*, vol. 381, no. 9868, pp. 752–762, 2013.

[29] World Health Organization, *Global Recommendations on Physical Activity for Health*, World Health Organization, Geneva, Switzerland, 2010.

[30] W. J. Chodzko-Zajko, D. N. Proctor, M. A. F. Singh et al., "Exercise and physical activity for older adults," *Medicine and Science in Sports and Exercise*, vol. 41, no. 7, pp. 1510–1530, 2009.

[31] K. Jin, J. W. Simpkins, X. Ji, M. Leis, and I. Stambler, "The critical need to promote research of aging and aging-related diseases to improve health and longevity of the elderly population," *Aging and Disease*, vol. 6, no. 1, pp. 1–5, 2015.

[32] G. B. J. Stainslow and J. Bempong, "Accessible smart home system for the deaf and hard-of-hearing," in *Proceedings of the New York Cyber Security and Engineering Technology Association (NYSETA) Fall Conference*, October 2015.

[33] C. S. Harrison, P. M. Grant, and B. A. Conway, "Enhancement of a virtual reality wheelchair simulator to include qualitative and quantitative performance metrics," *Assistive Technology*, vol. 22, no. 1, pp. 20–31, 2010.

[34] T. Perumal, A. R. Ramli, C. Y. Leong, S. Mansor, and K. Samsudin, "Interoperability for smart home environment using web services," *International Journal of Smart Home*, vol. 2, no. 4, pp. 1–16, 2008.

[35] T. Koskela and K. Väänänen-Vainio-Mattila, "Evolution towards smart home environments: empirical evaluation of three user interfaces," *Personal and Ubiquitous Computing*, vol. 8, no. 3–4, pp. 234–240, 2004.

[36] H. Tschofenig, J. Arkko, D. Thaler, and D. McPherson, "Architectural considerations in smart object networking," 2015.

[37] World Health Organization, *How to Use the ICF: A Practical Manual for Using the International Classification of Functioning, Disability and Health (ICF): Exposure Draft for Comment*, p. 5, WHO, Geneva, Switzerland, 2013.

[38] G. Stucki, A. Cieza, T. Ewert, N. Kostanjsek, S. Chatterji, and T. B. ÜstÜn, "Application of the International Classification of Functioning, Disability and Health (ICF) in clinical practice," *Disability and Rehabilitation*, vol. 24, no. 5, pp. 281–282, 2002.

[39] S. Arlati, D. Spoladore, S. Mottura et al., "Architecture of a virtual reality and semantics-based framework for the return to work of wheelchair users," in *Proceedings of the International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 74–85, Ugento, Italy, June 2017.

[40] A. Arcelus, M. H. Jones, R. Goubran, and F. Knoefel, "Integration of smart home technologies in a health monitoring system for the elderly," in *Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops, 2007 (AINAW'07)*, vol. 2, pp. 820–825, Niagara Falls, ON, Canada, June 2007.

[41] American Thoracic Society and others, "ATS/ACCP statement on cardiopulmonary exercise testing," *American Journal of Respiratory and Critical Care Medicine*, vol. 167, no. 2, p. 211, 2003.

[42] S. Kodama, K. Saito, S. Tanaka et al., "Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis," *JAMA*, vol. 301, no. 19, pp. 2024–2035, 2009.

[43] B. Matthews, "Semantic web technologies," *E-Learning*, vol. 6, no. 6, p. 8, 2005.

[44] N. Guarino, D. Oberle, and S. Staab, "What is an ontology?," in *Handbook on Ontologies*, pp. 1–17, Springer, Berlin, Heidelberg, 2009.

[45] M. Sacco, E. G. Caldarola, G. Modoni, and W. Terkaj, "Supporting the design of AAL through a SW integration framework: the D4All project," in *Proceedings of the International Conference on Universal Access in Human-Computer Interaction*, pp. 75–84, Heraklion, Crete, Greece, 2014.

[46] G. E. Modoni, M. Veniero, and M. Sacco, "Semantic knowledge management and integration services for AAL," in *Proceedings of the Italian Forum of Ambient Assisted Living (ForItAAL 2016)*, pp. 287–299, Pisa, Italy, 2016.

[47] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The NeOn methodology for ontology engineering," in *Ontology Engineering in a Networked World*, pp. 9–34, Springer, Berlin, Heidelberg, 2012.

[48] J. H. Gennari, M. A. Musen, R. W. Fergerson et al., "The evolution of Protégé: an environment for knowledge-based systems development," *International Journal of Human-Computer Studies*, vol. 58, no. 1, pp. 89–123, 2003.

[49] J. Z. Pan, "Resource description framework," in *Handbook on Ontologies*, pp. 71–90, Springer, Berlin, Heidelberg, 2009.

[50] S. Bechhofer, "OWL: web ontology language," in *Encyclopedia of Database Systems*, pp. 2008–2009, Springer, New York, NY, USA, 2009.

[51] I. Horrocks, P. F. Patel-Schneider, H. Boley et al., "SWRL: a semantic web rule language combining OWL and RuleML," W3C Member submission, vol. 21, p. 79, 2004.

[52] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, and Y. Katz, "Pellet: a practical owl-dl reasoner," *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 5, no. 2, pp. 51–53, Elsevier, 2007.

[53] S. Glass, G. B. Dwyer, and American College of Sports Medicine, *ACSM'S Metabolic Calculations Handbook*, Lippincott Williams & Wilkins, Indianapolis, IN, USA, 2007.

[54] F. J. P. Arts and H. Kuipers, "The relation between power output, oxygen uptake and heart rate in male athletes," *International Journal of Sports Medicine*, vol. 15, no. 5, pp. 228–231, 1994.

[55] C. E. Garber, B. Blissmer, M. R. Deschenes et al., "American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise," *Medicine and Science in Sports and Exercise*, vol. 43, no. 7, pp. 1334–1359, 2011.

[56] M. Peruzzini and M. Germani, "A Service-oriented architecture for ambient-assisted living," in *Proceedings of the 22nd ISPE Inc. International Conference on Concurrent*

Engineering (ISPE CE 2015), pp. 523–532, Delft, Netherlands, 2015.

[57] A. Goldfain, B. Smith, S. Arabandi, M. Brochhausen, and W. R. Hogan, "Vital sign ontology," *Bio-Ontologies*, vol. 2011, 2011.

[58] G. V. Gkoutos, P. N. Schofield, and R. Hoehndorf, "The units ontology: a tool for integrating units of measurement in science," *Database*, vol. 2012, p. bas033, 2012.

[59] American Society of Heating Refrigerating and Air Conditioning Engineers, *ASHRAE STANDARD: An American Standard: Thermal Environmental Conditions for Human Occupancy*, American Society of Heating Refrigerationg and Air Conditioning Engineers, Inc., Pennsylvania State University, State College, PA, USA, 1992.

[60] E. UNI 12464-1, "Illuminazione dei posti di lavoro," 2004.

[61] E. UNI 13779, *Ventilazione degli edifici non residenziali-Requisiti di prestazione per i sistemi di ventilazione e di climatizzazione*, Ente Nazionale Italiano di Unificazione, Milano, Italy, 2008.

[62] D. Spoladore, "Ontology-based decision support systems for health data management to support collaboration in ambient assisted living and work reintegration," in *Proceedings of the Working Conference on Virtual Enterprises (PRO-VE 2017)*, pp. 33–40, Cham, Switzerland, 2017.

[63] J. E. Crews and V. A. Campbell, "Vision impairment and hearing loss among community-dwelling older Americans: implications for health and functioning," *American Journal of Public Health*, vol. 94, no. 5, pp. 823–829, 2004.

[64] N. Fairhall, C. Sherrington, S. E. Kurrle, S. R. Lord, and I. D. Cameron, "ICF participation restriction is common in frail, community-dwelling older people: an observational cross-sectional study," *Physiotherapy*, vol. 97, no. 1, pp. 26–32, 2011.

[65] W. J. Rejeski, E. H. Ip, A. P. Marsh, M. E. Miller, and D. F. Farmer, "Measuring disability in older adults: the International Classification System of Functioning, Disability and Health (ICF) framework," *Geriatrics and Gerontology International*, vol. 8, no. 1, pp. 48–54, 2008.

[66] S. Nichols and H. Patel, "Health and safety implications of virtual reality: a review of empirical evidence," *Applied Ergonomics*, vol. 33, no. 3, pp. 251–271, 2002.

[67] D. A. Bowman, J. L. Gabbard, and D. Hix, "A survey of usability evaluation in virtual environments: classification and comparison of methods," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 4, pp. 404–424, 2002.

[68] D. Hix and J. L. Gabbard, "Usability engineering of virtual environments," in *Handbook of Virtual Environments: Design, Implementation and Applications*, vol. 28, CRC Press-Taylor & Francis, Boca Raton, FL, USA, 2002.

[69] J. Brooke, "SUS-a quick and dirty usability scale," *Usability Evaluation in Industry*, vol. 189, no. 194, pp. 4–7, 1996.

[70] G. Borg, *Borg's Perceived Exertion and Pain Scales*, Human Kinetics, Champaign, IL, USA, 1998.

[71] D. D. Price, P. A. McGrath, A. Rafii, and B. Buckingham, "The validation of visual analogue scales as ratio scale measures for chronic and experimental pain," *Pain*, vol. 17, no. 1, pp. 45–56, 1983.

[72] B. Dicicco-Bloom and B. F. Crabtree, "The qualitative research interview," *Medical Education*, vol. 40, no. 4, pp. 314–321, 2006.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 9302328, 12 pages
https://doi.org/10.1155/2017/9302328

Research Article

Helping Elderly Users Report Pain Levels: A Study of User Experience with Mobile and Wearable Interfaces

Iyubanit Rodríguez,[1] Gabriela Cajamarca,[1] Valeria Herskovic,[1] Carolina Fuentes,[2] and Mauricio Campos[3]

[1]*Department of Computer Science, Pontificia Universidad Católica de Chile, Santiago, Chile*
[2]*School of Computer Science, University of Nottingham, Nottingham, UK*
[3]*Medicine School, Pontificia Universidad Católica de Chile, Santiago, Chile*

Correspondence should be addressed to Valeria Herskovic; vherskovic@uc.cl

Received 28 July 2017; Accepted 4 October 2017; Published 6 November 2017

Academic Editor: Pino Caballero-Gil

Pain is usually measured through patient reports during doctor visits, but it requires regular evaluation under real-life conditions to be resolved effectively. Over half of older adults suffer from pain. Chronic conditions such as this one may be monitored through technology; however, elderly users require technology to be specifically designed for them, because many have cognitive and physical limitations and lack digital skills. The purpose of this article is to study whether mobile or wearable devices are appropriate to self-report pain levels and to find which body position is more appropriate for elderly people to wear a device to self-report pain. We implemented three prototypes and conducted two phases of evaluation. We found that users preferred the wearable device over the mobile application and that a wearable to self-report pain should be designed specifically for this purpose. Regarding the placement of the wearable, we found that there was no preferred position overall, although the neck position received the most positive feedback. We believe that the possibility of creating a wearable device that may be placed in different positions may be the best solution to satisfy users' individual preferences.

1. Introduction

Patient monitoring collects health information in real time [1], which can help health professionals improve treatment and diagnosis [2], while reducing health costs [3]. Chronic pain is described as "ongoing or recurrent pain, lasting beyond the usual course of acute illness or injury or more than 3 to 6 months, and which adversely affects the individual's well-being" [4]. Pain is frequent in the older population: up to 53% of elderly adults report suffering from recent pain [5]. Its treatment requires regular evaluation under real-life conditions in order to be resolved effectively [6]. Pain usually is measured through patients' self-reports only during medical appointments [7], using pain scales such as the Numerical Rating Scale or Verbal Rating Scale [8]. Pain is usually registered on paper, which can cause information loss and difficulties in analyzing and searching for data [9].

Users may be asked to report information remotely, for example, through the experience-sampling method, which asks participants to provide details about their current circumstances at certain intervals [10]. There have been several studies to explore whether elderly people self-report accurate information, with variable findings (e.g., [11, 12]). For patients with pain, self-reporting is believed to help patients become more aware of the characteristics of their pain, for example, its intensity, patterns, triggers, and location [13] and be more engaged in the self-management activity [14].

A wearable device is a "computer that is always with you, is comfortable and easy to keep and use, and is as unobtrusive as clothing" [15]. A wearable device must meet the following criteria: (1) the device is attached to the body and the user does not need to hold it, (2) the user does not remove the device to perform tasks or actions, and (3) the user must not separate the device from the body to interact with it [16]. In addition,

a wearable is on and working at all times [15]. Wearables can assist in monitoring patients with chronic pain during their daily routine, helping them better understand their illness and detect complications [17]. However, there is evidence of unequal access to technology, and, in some users, a lack of digital skills may hinder the possibility of using electronic devices for health monitoring [18]. The concept of *wearability* is used to describe how the wearable device interacts with the body [19] in terms of physical, emotional, and social comfort [20]. Designers of wearable devices must have knowledge of human physiology, since discomfort becomes evident when clothing impedes or restricts movement or visibility [21]. The *placement* of the device is one of the key concepts of wearability. The physiological, biomechanical, and comfort consequences should be included when evaluating a wearable device [16].

In this paper, we aim to explore the following research question: *are wearable interfaces appropriate devices for elderly people to self-report their pain level?*. To answer this question, we divided our research into two phases: the first aimed at understanding whether mobile devices or wearable devices are more appropriate for self-reporting pain levels and the second aimed at understanding which body position is more appropriate for elderly people to wear a device to self-report pain.

This paper is organized as follows. First, we discuss related work, especially focusing on existing technologies to report pain and studies on where wearable devices should be placed on the body. Then, Sections 3 and 4 describe each phase of the research, including the implemented prototypes, experiments, results, and discussion. Finally, Section 5 presents our conclusions and discusses limitations and possible avenues of future work.

2. Related Work

2.1. Technology and Interfaces for Elderly Adults. Aging is a process that depends on genetics, lifestyle, health [23], and gender, so the age in which a person is considered to be *elderly* varies. In several areas of the world, people over 50 are considered to be older adults [24], while the United Nations considers those over 60 to be older adults [25]. For the purposes of this study, and considering the cultural context of our study (conducted in Santiago, Chile), we consider people over 60 years of age and who are retired as elderly/older adults.

Elderly users require technology to be specifically designed for them, because many have cognitive and physical limitations [26], as well as a lack of digital skills that limits their ability to use electronic devices for disease monitoring [18]. Specifically in Chile, a high percentage of older adults have little experience with technology, almost 60% of 55- to 65-year olds in Chile have no computer experience whatsoever, while the average in OECD countries was 32% [27]. Regarding wearable devices, elderly users fear this type of technology may increase isolation and express concerns about safety and high costs [28]. Most studies with wearables and interaction have been conducted in countries in which older adults have a higher rate of digital skills (and a previous interest in technology is key in acceptance of wearables [29]), so our particular context provides additional challenges.

Several researchers have highlighted the importance of creating wearable technology for elders, for example, proposing six relevant considerations: motor, vision, eyeglasses, hearing, executive function, and memory [30]. Researchers have evaluated the use of commercial wearables by elderly adults (e.g., [31, 32]), finding them to be acceptable but require training for use. Less mainstream wearables (e.g., head-mounted displays) have been more complex for elders [33]. It is somewhat clear that established wearables such as activity trackers, that require little interaction from users, result in a better user experience than more experimental, niche wearables or than wearables that require users to interact with them.

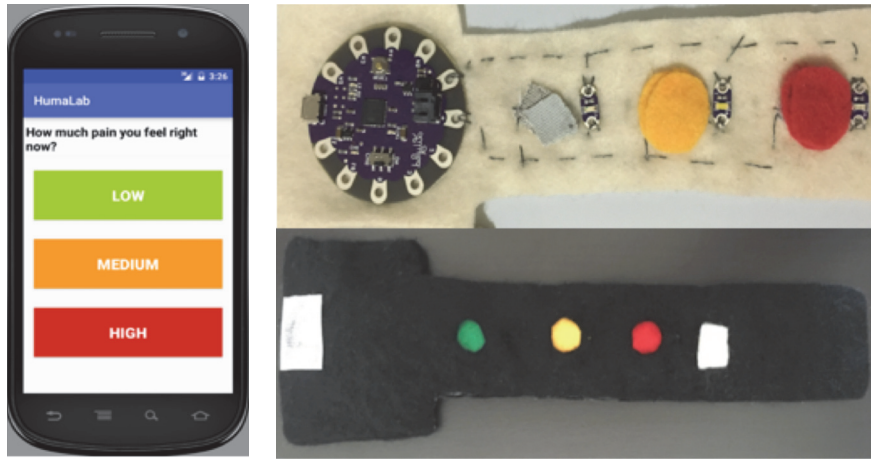
2.2. Technology for the Self-Report of Pain. A system for patients to report pain from anywhere, at any time, can be used to monitor the evolution of pain [6]. Several applications allow people to report episodes of pain; for example, one mobile application displays a human figure and asks the user to indicate the position, intensity, and type of pain he/she feels [34, 35] or a web application that uses a combination of body diagrams and a Numerical Rating Scale, providing the transfer of patient information to health professionals [36].

Researchers have proposed several novel interfaces to self-report pain. A tangible device that allows users to easily record their pain using a six-level scale found that these types of pain-recording devices may decrease the pain experience [6]. Wearables with interactive displays allow users to input information [37]. For example, a wearable device to self-report pain and emotional state found that it may help users improve their self-knowledge [38]. These types of applications allow new avenues of patient-doctor interaction [9], but adherence rates are often low [39] because some are burdensome [14], or not portable.

2.3. Placement of Wearable Devices on Body. The placement of wearable devices for elderly people is an important issue. A wearable device must allow easy access and handling [40], be discreet, ergonomic, and well affixed [17], and allow body movement [21] and visibility [41]. Placement should be in areas that are relatively the same size across adults (with a large surface area) and with low movement when the body is moving [19].

Regarding the best location of these devices in the body, eight possible areas for the unobtrusive placement of wearables have been identified, for example, collar area, rear of the upper arm, waist and hips, thigh, and top of the foot [19]. In the case of biomedical sensors and devices, they have been placed on headbands, helmets, belts, shoes, socks, bracelets, arm bands, and shirts [17]. Furthermore, large and curvilinear areas can be used in skin interfaces, for example, the back, the back of the hand, and the neck [42].

A recent study compared the placement of a wearable for elders on the wrist, upper arm, and neck, finding that the wrist was the best location (allowing the best viewing angle, a greater willingness to exhibit the device, and less anxiety) but that personal characteristics affected preferences [41]. Another study found that the wrist has the advantage that it is positioned approximately in the same place and orientation



(a) (b)

Figure 1: (a) *PainApp*: mobile app. (b) *B-pain*: wearable device.

for all users [43]. Previous studies have not focused on where to place devices when elderly people have to interact with the device (through reporting information), which is the focus of this work.

3. Phase I: Comparing Wearable Devices and Mobile Applications

The first phase of this research focused on comparing two prototypes (one mobile and one wearable) that allow elders to self-report pain during their daily lives. A detailed account of the implemented devices and the experiment conducted in this phase can be found in [44]. Both of the devices use a simple Verbal Rating Scale with three levels of intensity (Low, Medium, and High). The two prototypes are described as follows:

(i) *PainApp*: it is a simple Android application that asks users about their pain level and stores the information in a database. The information can then be shared through email, Bluetooth, or social networks (Figure 1(a)).

(ii) *B-pain*: it is a bracelet-shaped wearable device, implemented using LilyPad Arduino. Pain is reported by pressing one of three buttons (green = low; yellow = medium; red = high). The device provides feedback through a LED light (Figure 1(b)).

3.1. Materials, Methods, and Participants. The evaluation was done during May 2016. The participants were 12 undergraduate students (6 women and 6 men). The average age of participants was 26 (standard deviation = 5.4). All participants had *above basic* digital skills. We applied semistructured interviews (each interview lasted about 15 minutes). To evaluate the solutions the participants interacted with the mobile application and wearable device and then we collected four types of information:

(i) *DIGCOMP*: DIGCOMP is a standardized instrument to measure digital competences, where users are categorized into one of four possible groups, according to their digital skill levels: *none*, *low*, *basic*, or *above basic* [45].

(ii) *System usability scale (SUS)*: SUS is a quick way to measure the overall usability of the system [46]. In this scale, scores below 60 indicate poor usability, while scores over 80 indicate very good usability [47].

(iii) *Usability questionnaire*: a questionnaire regarding usability of the wearable device.

(iv) *Interview data*: interviews were recorded (audio) and transcribed. Subsequently, each interviewee was assigned a code (P1 to P12).

3.2. Experiment. To avoid bias, half of the participants interacted with the mobile application first and the wearable device second, and the other half performed the opposite process. Each interview had the following structure:

(1) One researcher gave a brief introduction about the study and its purpose (5 minutes).

(2) The participant read a scenario describing a person with pain (a college student who after a car accident is suffering from chronic back pain) (5 minutes).

(3) The first interface was explained (3 minutes).

(4) The participant was given time to interact with the first interface (3 minutes).

(5) The researcher asked questions about the interface, using a predefined question set to guide the discussion (8 minutes).

(6) Steps (3)–(5) were repeated with the second interface (14 minutes).

(7) The researcher asked questions comparing the two interfaces (5 minutes).

(8) Finally, participants completed the DIGCOMP, SUS, and user experience questionnaires (8 minutes).

3.3. Results. The interview data was transcribed, and thematic analysis was used for codification and analysis [48]. Some quotes from participants are provided in the results (translated from Spanish). We aimed to answer two questions: (1) *which is more appropriate (has a higher rate of user acceptance) for monitoring pain: a wearable interface or a mobile application?* And (2) *which characteristics, or features, of the wearable interface, are critical for users to be able to report their pain levels?*.

Regarding simplicity, 67% of participants found the wearable device (B-pain) was simpler, while only 16.5% found the mobile application simpler (and 16.5% found that both technologies are equally simple). P4 said the following: *"the bracelet, because it's easy and fast. If I feel pain I just need to push a button, while with the cellphone I have to turn it on, open the application, and then report pain."* The wearable device was found by 75% of users to report pain at the right time, whereas in the mobile application the user was delayed by opening his/her phone and finding the app. Regarding the digital skills necessary to use each interface, 41.6% of participants believed the mobile application required some knowledge about how to use a smartphone, while 33% believed the wearable device only required a brief initial explanation about how to use it.

The key features of the wearable interface that are needed for users to be able to report their pain are the following ones:

(i) Low cognitive load: a device should require a low cognitive load and be simple to understand. For instance, B-pain has only one functionality. One user did mention that this device might not work for colorblind users, so it is important to complement the interface with, for example, words or textures that can help users with disabilities or other conditions use them.

(ii) Anytime/anywhere availability: a device to self-report pain should be easily available at all times, facilitating access to self-report. We call this anytime-anywhere availability *"when you need it, you have it" (WYNIYHI)*.

(iii) Materials: it is important to consider the type of material with which the wearable device is designed. Materials can cause problems for users, for example, allergies.

(iv) Pain intensity and feedback: users should have the possibility of reporting several pain intensity levels (e.g., a 10-point numerical scale), but without making the design more complex. The feedback to the user should be clear and at the right time, so that the user is aware that the actions have been properly completed.

3.4. Discussion. Overall, this study found that both mobile and wearable applications may be appropriate for users to self-report pain, depending on the users and their context. For instance, in some situations, a wearable device may interfere with clothing conventions, and a mobile application (installed in a phone that the user would be carrying anyway) would be more appropriate. As evidence that pain-reporting applications are appropriate, a previous study found 25 diary or journal-type applications relating to pain, in which users could register their pain levels and optionally share data with health professionals and researchers [49]. However, in the current study, the wearable device was more widely accepted for monitoring pain than the mobile device.

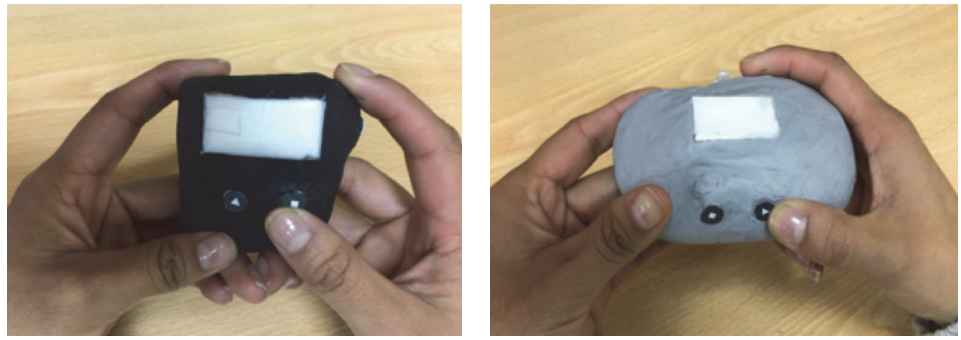
As in previous studies, for example [50, 51], the fact that the wearable device had a clear purpose affected its acceptance. Although a wearable device to report pain would mean carrying an additional device, the surveyed users did not feel this was a limitation. The participants liked the immediate accessibility and limited functionality of the wearable device. One possible explanation is that users perceive mobile phones as a tool of social communication, while wearables are perceived as well-being devices [51].

The acceptance of a wearable device is affected by several social factors [52], among them perception of ease of use and mobility. Regarding ease of use, if a wearable device aimed at healthcare/well-being is perceived to be complicated, this could generate anxiety because a user might think that a mistake could be harmful to their health [52]. In the case of *B-pain*, the device is extremely simple and intuitive, serving only one function and with limited user interaction, which results in high ease of use. This ease of use has been found to give the user a feeling of control over the device [53]. Regarding mobility, although cellphones are highly personal and usually close to their user, *B-pain* is a bracelet, so it can also be carried anywhere easily. Since *B-pain* is worn on the wrist, similar to a watch or bracelet, it does not impede movement [19].

Previous research has found that portable devices should be lightweight, durable, and comfortable and with good appearance [51]; this was also evident in our study, as people were very emphatic in the importance of improving materials and aesthetics. Another aspect to consider in the appearance of wearable devices is the surrounding context and culture in which they are used, for example, the material, colors, and type of clothing may vary depending on the activity [21], so a wearable should be able to adapt to different contexts. The participants in our experiment mentioned that they would use both the wearable and the mobile application, since the phone may be more convenient for social situations, because it does not interfere aesthetically with the user's clothing.

This work found a trade-off between the usability of the wearable and the difficulty in replicating and maintaining it, since the mobile application is easily replicable and easily disseminated, while the portable device, although easier to use and requiring a lower cognitive load, requires hardware and must be built.

This phase of evaluation has several limitations that we would like to acknowledge. First, the participants were all students, so the digital skills and context are different from elders, and it is not possible to ascertain whether the results would have been the same if elders had been interviewed. The context of this study was a higher-education institution in Santiago, Chile, which could also account for cultural factors. Second, the participants did not use the device/application



(a) (b)

Figure 2: (a) Model A (7 × 5 cm). (b) Model B (6 × 9 cm).

during their daily life, rather they evaluated the interfaces after a brief period of interaction. Third, the participants were not users with pain, rather they were asked to imagine a context in which they were suffering from chronic pain and had to report it. Fourth, the size of the study was small, so we mainly focused on qualitative data. Although all of these limitations threaten the generalizability of our study, we found this first phase provided valuable insights, identifying concerns about how to design a wearable device that may allow elders to self-report pain, and we believe this is an important step before interviewing the actual targets of the device (elderly users).

4. Phase II: RepWear (A Wearable to Self-Report Pain)

Phase I found that a wearable device for self-report of pain is preferable over a mobile application, due to its simplicity, limited functionality, and greater accessibility. Therefore, since our goal was to design a device to self-report pain for elderly users, we set out to design a new prototype that could be used and tested by elderly users. In this phase, the focus was on the placement of a wearable device for elderly people to self-report their pain levels.

The design of this wearable device was based on several guiding principles: first, interfaces for elders should be simple and not use excessive instructions [54]. Second, the design should be based on a familiar concept to older adults [55]. For these reasons, and based on the previous study, we developed a simple, easy to use, intuitive prototype that requires little cognitive effort.

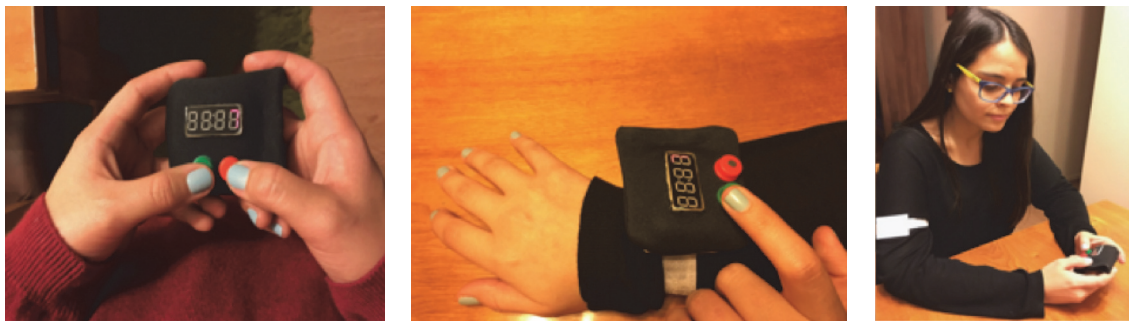
The wearable device we designed, called *RepWear*, allows self-reporting of pain by using two buttons: one allows navigating the intensity of pain from 0 to 10, and the other one stores the selected value. In addition, a display shows the values of pain intensity and the device has an on/off switch. Two nonfunctional prototypes were created in order to understand people's perceptions about the preferred physical characteristics of such a device. Figure 2 shows two models that were made with modeling material and plastic 3D-printed pieces to simulate Arduino boards and buttons. We showed the prototypes to 10 adult participants (8 women, 2 men, average age: 37.2, SD: 19.98) and asked them to fill out a SUS questionnaire and participate in a brief interview. Nine out of 10 participants preferred *Model A* to *Model B*, stating that *"this one is smaller, if it's big it scares me, simpler is better for an old man [like me]…."* The results also showed a greater acceptability of Model A, since it received a SUS score of 90, while Model B had a score of 77. The participants gave reasons for their selection such as the location of the buttons, the position of the switch, and the size, while recommending that (1) the buttons must have different colors, while keeping the icons and (2) the shape should be more curved. Finally, the participants mentioned that the positions where the device could be placed on the body were wrist, neck, waist or belt, arm, and pants pocket or/and shirt pocket.

Therefore, based on the previous results, a new, functional prototype was designed, incorporating a Numerical Rating Scale (0 to 10), anytime/anywhere availability, softer material, feedback through sound and touch, and a curved shape with colored buttons. The device was implemented using a LilyPad Arduino Main Board (electronic card based on ATmega168V) and a Lithium Ion Battery, 1 Ah. A Real-Time Clock (RTC) (DS1307) was used to record the time and date when the user self-reports pain intensity. The device incorporated two buttons: a red one to allow users to save pain intensity and a green one to allow users to select a pain intensity. The user may see the reported number on the display (Grove, 4-Digit Display module). When users self-report pain, they receive feedback because the button is mechanical and clicks. Figure 3 shows a user self-reporting pain. The information (time and date, pain intensity) is saved in a microSD card. A miniature DPDT slide switch may be used to turn the device on or off.

4.1. Materials, Methods, and Participants. The evaluation was done during June and July 2017. The participants were 18 elderly people (13 women and 5 men). The participants were all older adults, ranging in age from 60 to 93 (average: 69.3, SD: 9.36). They did not have any mental disabilities. Table 1 describes each study participant. Participants with *None* digital skills mentioned that they only use the phone to call and/or send messages. In addition, two of the participants (P4 and P5) walk with support: one with a walker and the

Table 1: Description of study participants.

P	Age	Gender	Occupation	Educational level	Digital skills	Location of residence	Bedridden	Pain location
P1	60	F	Housewife	High school	None	Own house		Knee
P2	82	F	Housewife	High school	None	Nursing home	•	Back
P3	93	M	Mechanic	Primary	None	Nursing home	•	Prostate
P4	73	M	Teacher	Master	Above basic	Nursing home		Hands and legs
P5	80	F	Housewife	Primary	None	Nursing home		Legs
P6	70	F	Teacher	University	Basic	Own house		Back
P7	65	F	Paramedic	Technical	Basic	Own house		Back
P8	60	F	Therapist	Technical	Above basic	Own house		Back
P9	62	M	Grocer	High school	Basic	Own house		Back
P10	76	F	Housewife	High school	None	Own house		Hip
P11	79	M	Building	None	None	Own house		Neck
P12	60	M	Teacher	University	Basic	Own house		Back
P13	60	F	Secretary	Technical	Basic	Own house		Back and foot
P14	62	F	Housewife	High school	None	Own house		Back and foot
P15	65	F	Housewife	School	None	Own house		Breast
P16	72	F	Secretary	Technical	None	Own house		Back and stomach
P17	68	F	Housewife	School	None	Own house		Column and knee
P18	60	F	Housewife	High School	None	Own house		Neck and hand



(a) (b) (c)

Figure 3: (a) *RepWear* prototype (6 × 6.3 cm). (b) Self-report pain from the device on the wrist. (c) Extracted from the strap to self-report.

other with the help of another person. Two participants (P2 and P3) were bedridden.

To evaluate the placement of the device, the participants interacted with the device and placed it in 4 different body positions: wrist, waist, neck, and arm. The following four types of information were collected:

(i) *DIGCOMP*: DIGCOMP is digital skills questionnaire (see Section 3.1).

(ii) *Wearable satisfaction questionnaire*: we created a 7-point Likert scale questionnaire based on the findings of a previous study on placement of wearables for elderly people [41]. We created one question for each of the following issues: *willingness to show the wearable device*, *anxiety*, *oddness*, *fear of others' negative reaction*, *comfort*, *readability of the device screen*, and *unobtrusiveness for daily activities*.

(iii) *AttrakDiff questionnaire*: AttrakDiff is a questionnaire used to understand how users personally rate the usability and design of an interactive product. AttrakDiff has four dimensions: *pragmatic quality*, which is the ease with which the user can complete the task, *hedonic quality-identity (HQ-I)*, which is the message that is communicated to others while the product is being used, *hedonic quality-stimulation (HQ-S)*, which represents whether the development of user skills is encouraged, and *attraction*, or the overall charm of the product. Answers are on a scale of −3 to 3 (0 represents neutrality) [22, 56].

(iv) *Interview data*: each interview was recorded (audio), transcribed, and assigned a code (P1 to P18).

4.2. Experiment. To understand the perceptions of older adults about the placement of a wearable device to self-report pain, we chose 4 body parts to investigate: the wrist, arm,

Figure 4: Use of *RepWear* while doing activities of daily living in the four body positions: wrist, arm, waist, and neck, respectively.

waist, and neck. The experiment had a duration of between 60 and 75 minutes per participant. The activities carried out were the following ones:

(i) The researcher explained the purpose of the investigation and answered questions about it, and then the participant signed the informed consent form (10 to 15 minutes).

(ii) The participant was asked basic information such as age, profession, education level, and pain location. They completed the DIGCOMP questionnaire regarding digital skills (5 minutes).

(iii) The participant used the device to self-report pain at each position of the body for 6 minutes (24 minutes for all four positions), while they performed their everyday routines and tasks (this process is shown in Figure 4). The researcher observed and took notes. Participants were asked to self-report pain once for each position, either directly from the device or by removing it from its strap and then returning it to the strap (see Figures 3(b) and 3(c)).

(iv) After completing each position, the participant answered the wearable satisfaction questionnaire.

(v) After completing all the positions, the researcher conducted a semistructured interview to collect information from participants about the preferences of the body positions to carry the device (ranking) and their perception about the device (5 to 10 minutes).

(vi) Finally, the participant completed the AttrakDiff questionnaire about *RepWear* (10 to 15 minutes).

4.3. Results

4.3.1. Body Placement Results. Participants were asked to rank their preferred placements for the device. One of the participants chose not to provide a ranking, as he/she felt all the positions were approximately the same. The ranking data was analyzed using the Shapiro-Wilk test, obtaining that the distribution is not normal ($\alpha = 0.05$). Then, the Kruskal-Wallis test was applied ($n = 17$), with the result that the ranking of positions (neck, arm, waist, and wrist) did not have a statistically significant difference overall. However, we analyzed the qualitative responses in depth to understand the user experience with each possible body position. Next, we provide the analysis of each position. Quotes from the interviews are provided, translated from Spanish.

(i) *Neck*: the neck position was comfortable and loose for the participants *"because... let's see, it did not bother me at all, really at all. It's like wearing a necklace: more comfortable, friendlier"*. Also, the readability of the screen was good: *"it's comfortable, it's not bothersome, because from here I can see [the pain scale]."*

(ii) *Wrist*: responses about this position were polarized. Some felt that wearing the device on the wrist interfered with their activities: *"it is closer to the hand with which I do things... it could bother me when grabbing things or separating [papers]. It would disturb me more"*, or were worried about it falling or becoming damaged *"it may inconvenience me more, because of where it is placed, I could hit it on something"*, *"it is uncomfortable because we are moving our hands at all times and it may fall. I can do things but it may fall. That's the danger I see."* Other participants gave positive comments to the position of the wrist as being adequate to see the pain scale: *"it's more comfortable to manipulate the device and look at it [the pain scale]"* and for comfort *"...more comfortable to move... for mobility."*

(iii) *Arm*: this position had very few comments from the participants. People who ranked it first thought that it was comfortable and those who ranked it last thought the opposite: *"because of my usual activities, it would be uncomfortable: to throw a ball, to jump..."*. Users also thought they would have mobility problems *"it's more uncomfortable... it would not allow me to move"*.

(iv) *Waist*: the waist was the position that had the most negative comments. Participants emphasized that readability is low, since in this position the device could be covered by clothes: *"I always wear something*

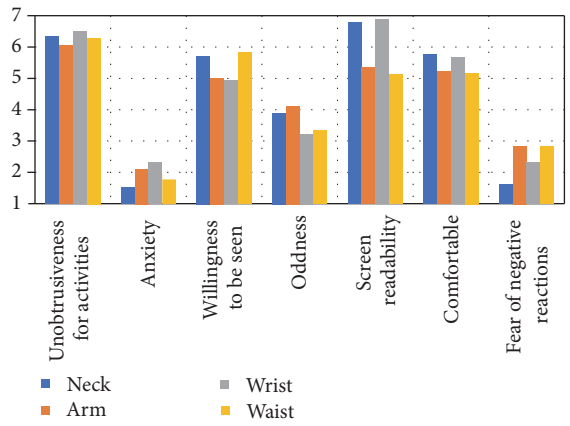


Figure 5: The average satisfaction degree in four parts of the body (units = points, 1 (strongly disagree) to 7 (strongly agree)).

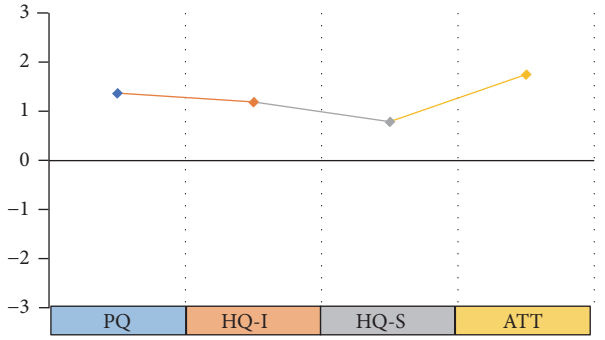


Figure 6: Average values for pragmatic quality (PQ), hedonic quality-identification (HQ-I), hedonic quality-stimulation (HQ-S), and attractiveness (ATT) (results from [22]).

> *that covers my waist. Even though I am thin I feel more secure with my waist hidden from view and compressed."* Another participant indicated that the waist position did not adjust to heavier shapes: *"I am bigger and fat."* However, three people mentioned that the waist is comfortable to report the pain and that the device will be in less danger: *"I can put it in the front and it will be more protected, I'll take better care of it."*

The results from the wearable satisfaction questionnaire (Figure 5) for each position of the body regarding unobtrusiveness, anxiety, willingness to be seen, oddness, screen readability, comfort, and fear of negative reactions show that participants mostly gave similar scores to each position and that the results were positive. For example, users felt that the device was unobtrusive (average score: 6.29) and caused low anxiety (average score: 1.93), they were willing to be seen (average score: 5.83), and it was more or less odd (average score: 3.64), readable (average score: 6.03), and comfortable (average score: 5.46) and did not cause much fear of negative reactions (average score: 2.40).

These responses were analyzed using the Shapiro-Wilk test ($\alpha = 0.05$), finding that the responses did not follow a normal distribution. Then, the Kruskal-Wallis test was performed ($n = 18$, $\alpha = 0.05$), finding that the only statistically significant difference was the readability of the device (p value: 0.0358). Then, we applied the Mann–Whitney test, finding that the neck and wrist had significantly higher scores than the waist for readability (p values: 0.0227 and 0.004).

We discuss each of the items as follows:

(i) *Willingness to show the wearable device*: participants said that they would show the device in public only if it was absolutely necessary: *"I would use it only if I felt a lot of pain."* One participant mentioned that wearing the device on the waist may help the device be unnoticed.

(ii) *Fear of others' negative reaction*: all the participants felt little fear of the reaction of others when carrying the device: *"at this age I don't care what others think of me, I am too old for that [laughing]."* Also, three participants commented that they would fear that the device could be stolen.

(iii) *Readability of the device screen*: the positions that offer a better view of the pain scale are the wrist and neck with a score of 6.89 and 6.78, respectively. This is because in those positions people can self-report pain without removing the device from the strap, whereas in the arm (5.33) and the waist (5.11) the majority of the participants (13 people) chose to remove the device from strap to (1) have a better view of the scale and (2) be able to handle the button.

(iv) *Anxiety*: the reported scores for anxiety were very low; that is, the participants did not feel anxious when using the device.

(v) *Oddness*: participants did not feel odd when using the device, rather they felt that others might: *"people could feel strange when they see a device like this. . .."* The participants felt that wearing the device on the wrist was not odd because of the similarity of the device with a wristwatch: *"for me, it's like wearing a watch, for that reason I liked it more, because I am used to it."*

(vi) *Comfortable*: the majority of the participants were comfortable using the device in the 4 parts of the body; however, the neck (5.78) and the wrist (5.67) received the highest scores, *"I feel I can wear it comfortably and securely."*

(vii) *Unobtrusiveness for daily activities*: participants used the device to perform their daily activities to determine which position was most suitable for that purpose. The results showed that when the device was placed on the wrist (6.50), there were fewer interruptions or discomfort.

4.3.2. User Experience Results. User experience (UX) is defined as a "person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service" [57], including aspects of product use, reflections of interactions, user expectations, and feelings [58]. User experience was measured through the AttrakDiff questionnaire (Figure 6). RepWear was found to have a high pragmatic

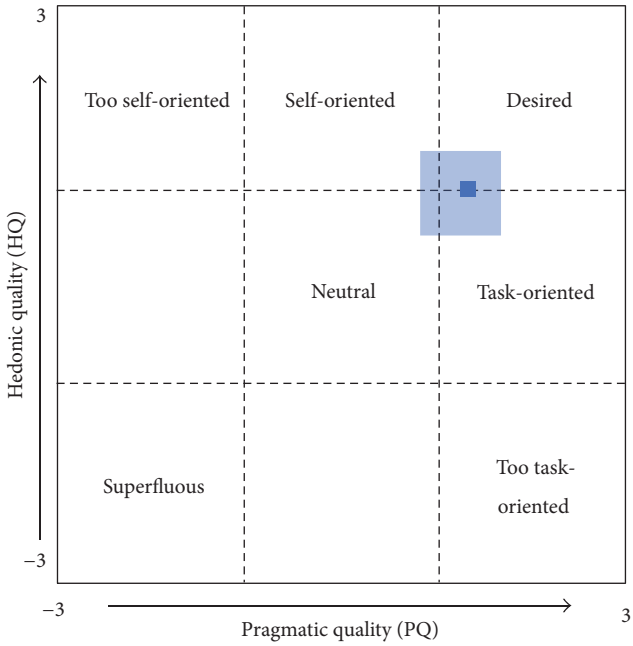


Figure 7: Portfolio with average values of the dimensions PQ and HQ and the respective confidence rectangle (results from [22]).

quality (PQ = 1.37), which means users can do the task of self-reporting pain. Nevertheless, hedonic quality was low (HQ = 0.99) leaving the device in between the task-oriented and desired categories (Figure 7). The confidence levels were 0.46 for PQ and 0.48 for HQ, which means that answers were a little scattered. The attractiveness dimension (ATT) had a score of 1.75, which means that participants had a positive assessment of the appeal of the device. The score for hedonic quality-identification (HQ-I), or "how others can see the product," was 1.19. The hedonic quality-stimulation (HQ-S) score was the lowest (0.79), meaning that RepWear was not perceived as a device that helps to develop personal skills and improve knowledge. Figure 8 shows that the word pairs *undemanding-challenging* and *cautious-bold*, pertaining to the HQ-S, had scores below 0.

During the interviews, participants were asked their opinions about the device. They mentioned that RepWear was easy to use, handy, and useful: *"I found it was novel, and good in the sense that one can have it on all day to mark pain intensity and the time it happened."* Also, they thought the pain scale was appropriate: *"it's good that it asks me on a scale from 0 to 10... it's much better."* Another participant mentioned that the colors on the buttons reminded her how to use the device: *"... I liked the colors from these two little things, it's like calling on the phone."*

Regarding the aspects of RepWear that participants did not like, four mentioned that they would prefer a thinner, lighter device: *"it protrudes too much, it should be smaller, more adaptable to the body."* Nine participants said there was nothing that they did not like about the device. Additionally, one participant expressed that he did not like the device at all because it was not useful.

One participant thought that the buttons were too sensitive and suggested improvements: *"the numbers, when I push it's too sensitive, it changes way too fast."* Another participant would like the device to be more firmly attached to the body: *"it should be fixed to a spot, because it's not secure... maybe a strap would make it more secure. This could fall."* Participants also commented on the device material and color, with some suggesting different colors or using plastic material.

4.4. Discussion. Our results suggest that the neck may be the best position for a wearable to self-report pain, since it generally had good scores in all of the wearable questionnaire items (especially regarding anxiety, fear of negative reactions, screen readability, and comfort). These results differ from previous results that suggested that a neck wearable device would be annoying and uncomfortable [41]. This may be because of cultural reasons or because RepWear required self-report, so this potential interaction may affect how users perceive the device. Further study is needed to examine these discrepancies and understand the reasons behind them.

Our device, although designed as a wearable, did not completely fulfill one of the criteria for wearables, that is, that the device does not need to be separated from the body for interaction. When the device was worn in the waist and arms, the participants chose to remove it to manipulate it, changing the device from a wearable to a portable device. This allows us to identify that placement is key: the position of the device at the body has an impact into facilitating or hindering visualization and self-report.

The AttrakDiff questionnaire measures hedonic qualities such as stimulation and innovation. In the case of an interface designed for elderly users with low digital skills, we wanted to create an interface that was not intimidating. In this sense, the low HQ-S results may be interpreted in a positive light, since they are related to adjectives such as *undemanding* and *cautious*, showing that the users did not feel threatened or challenged by the device. We believe adding buttons that clicked mechanically was partially the cause of this result. Previous studies have found that elders feel *computer anxiety* [59] about damaging a technological device while using it, anxiety about "doing something wrong." For example, elderly users are anxious when moving a computer mouse [60]. However, in our study, even elders with no digital skills did not feel anxious, except when using the device on the wrist, which they felt could damage the device (e.g., when washing dishes).

It is important to mention that the evaluated device was a prototype, lacking the more polished look and feel of a finished product, as well as being bulkier. These characteristics of the prototype could influence the perception of the participants when trying on the device. Generally, when designing devices such as this one, the thickness of the device should be minimized, allowing greater safety and comfort [19]. Our prototype had a thickness of 2.5 cm, which may cause some discomfort in the participant if it is worn on certain parts of the body such as the wrist. Also, elders tend to prefer a compact device for portability [41] and movement [52]. In this regard, the *RepWear* prototype still has room

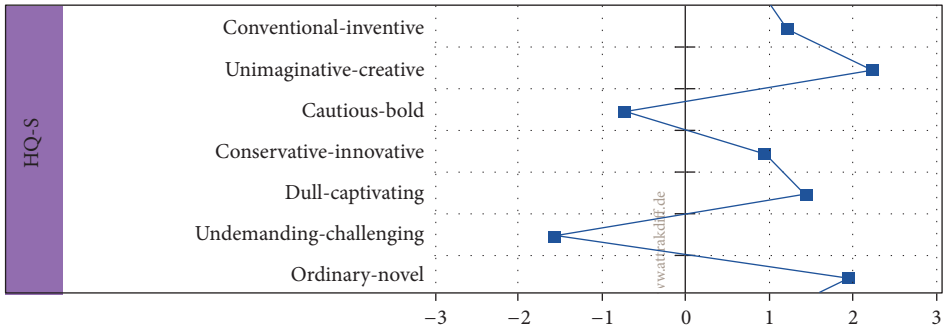


Figure 8: The mean values of the word pairs of hedonic quality-stimulation (HQ-S) (diagram from [22]).

for improvement and size, thickness, and weight should be further decreased.

In this phase the participants were elderly adults, which allows for some comparison with the first phase of our research. We found that the elderly users were less concerned with aesthetics; they did not find the device overly odd nor feared negative reactions, and they did not comment extensively on materials nor matching clothing as the participants in the first phase. Rather, the elderly participants mainly focused on the functionality and comfort of the device, especially in regard to being able to carry out their activities normally. The literature has also found perceived utility to be of great importance for elders, since they will use a system if they consider it to be useful and reliable and provide benefits to their independence [52]. In our study, only one participant (out of 18) did not find the device to be useful. Finally, the cost of a device can influence that user to use it, and an important challenge is for technology to be of low-cost [61]. Our device is of low-cost, using simple components (and would be extremely low-cost if it was mass-produced).

We would like to acknowledge the study limitations for this phase. First, the number of participants in the study was small, because of the difficulty in recruiting elderly participants, since we only chose to interview those without cognitive impairments, and the extensive questionnaires we used caused the participants some cognitive overloading, causing them to become tired. This may suggest the need for using questionnaires that are especially designed for the elderly, in order to improve their willingness to participate. Second, the study was conducted in Santiago, Chile, and cultural characteristics may make its results not generalizable to a broader region.

5. Conclusion and Future Work

The goal of this research was to explore the best way for older adults to report pain. To achieve this goal, we divided our research into two phases: (1) finding out whether a mobile or wearable interface would be more appropriate and (2) finding where to place the wearable device.

First, we found that users preferred the wearable device over the mobile application and that a wearable to self-report pain should be designed specifically for this purpose, be aesthetically pleasing, and allow users to report easily and at the right time. Second, we found that although participants had individual preferences for the placement of the wearable device, there was no preferred position overall. The neck position received the most positive feedback, because it produced less anxiety, there was no fear of the reaction of other people, and it was comfortable and had a good view of the screen. We believe that the possibility of creating a wearable device that may be placed in different positions may be the best solution to satisfy users' individual preferences.

We would like to acknowledge some of the study limitations. First, the study should be expanded to include a higher number of participants, although there are several challenges in recruiting older participants without cognitive impairment, and extensive questionnaires were found to cause some cognitive overloading. Second, the first phase of the study used student participants (due to the difficulties in recruiting elderly people), and these results may therefore not completely be applicable to the elderly. Third, the study was conducted in Santiago, Chile, and cultural characteristics may not make the results generalizable to a broader region.

As future work, we will analyze how to improve certain physical aspects of the device (e.g., decreasing size and thickness), in order to make the device more comfortable and versatile. We will also work on recruiting additional participants to explore in more depth the reasons behind their placement preferences.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This project was supported partially by CONICYT-PCHA/Doctorado Nacional, 2014-63140077, CONICIT and MICIT Costa Rica Ph.D. scholarship grant, Universidad de Costa Rica, and CONICYT/FONDECYT no. 1150365 (Chile). The authors thank Javiera Rivas, Leonor Toro, and Teresa Uribe for their help with the experiment and the nursing homes that opened their doors to them.

References

[1] World Health Organization, *mHealth: New Horizons for Health through Mobile Technologies*, vol. 3 of *Global Observatory for eHealth Series*, Geneva, Switzerland, 2011.

[2] A. Alahmadi and B. Soh, "A smart approach towards a mobile e-health monitoring system architecture," in *Proceedings of the 2011 International Conference on Research and Innovation in Information Systems, ICRIIS'11*, November 2011.

[3] M. J. Field and J. Grigsby, "Telemedicine and Remote Patient Monitoring," *Journal of the American Medical Association*, vol. 288, no. 4, pp. 423–425, 2002.

[4] ACPA., Glossary @ONLINE, 2016.

[5] K. V. Patel, J. M. Guralnik, E. J. Dansie, and D. C. Turk, "Prevalence and impact of pain among older adults in the United States: findings from the 2011 National Health and Aging Trends study," *PAIN*, vol. 154, no. 12, pp. 2649–2657, 2013.

[6] I. Alakärppä, J. Riekki, and R. Koukkula, "Pervasive pain monitoring system user experiences and adoption requirements in the hospital and home environments," in *Proceedings of the 2009 3rd International Conference on Pervasive Computing Technologies for Healthcare - Pervasive Health 2009, PCTHealth 2009*, pp. 1–8, April 2009.

[7] A. Chhikara, A. Rice, A. McGregor, and F. Bello, "In-House Monitoring of Low Back Pain Related Disability (IMPAIRED)," in *Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4507–4510, August 2008.

[8] A. Williamson and B. Hoggart, "Pain: a review of three commonly used pain rating scales," *Journal of Clinical Nursing*, vol. 14, no. 7, pp. 798–804, 2005.

[9] F. Spyridonis, J. Hansen, T.-M. Grønli, and G. Ghinea, "PainDroid: An android-based virtual reality application for pain assessment," *Multimedia Tools and Applications*, vol. 72, no. 1, pp. 191–206, 2014.

[10] R. Larson and M. Csikszentmihalyi, "The experience sampling method," *New Directions for Methodology of Social & Behavioral Science*, 1983.

[11] T. L. Bush, S. R. Miller, A. L. Golden, and W. E. Hale, "Self-report and medical record report agreement of selected medical conditions in the elderly," *American Journal of Public Health*, vol. 79, no. 11, pp. 1554–1556, 1989.

[12] K. Hilton, J. Fricke, and C. Unsworth, "A comparison of self-report versus observation of performance using the assessment of living skills and resources (ALSAR) with an older population," *The British Journal of Occupational Therapy*, vol. 64, no. 3, pp. 135–143, 2001.

[13] C. Rini, D. A. Williams, J. E. Broderick, and F. J. Keefe, "Meeting them where they are: Using the Internet to deliver behavioral medicine interventions for pain," *Translational Behavioral Medicine*, vol. 2, no. 1, pp. 82–92, 2012.

[14] H. MacLeod, A. Tang, and S. Carpendale, "Personal informatics in chronic illness management," in *Proceedings of the Graphics Interface GI '13*, pp. 149–156, Canadian Information Processing Society, Toronto, Canada, 2013.

[15] B. J. Rhodes, "Wearable remembrance agent: a system for augmented memory," in *Proceedings of the 1997 1st International Symposium on Wearable Computers*, pp. 123–128, October 1997.

[16] J. F. Knight, D. Deen-Williams, T. N. Arvanitis et al., "Assessing the wearability of wearable computers," in *Proceedings of the 10th IEEE International Symposium on Wearable Computers, ISWC 2006*, pp. 75–82, October 2006.

[17] A. Dittmar, R. Meffre, F. De Oliveira, C. Gehin, and G. Delhomme, "Wearable medical devices using textile and flexible technologies for ambulatory monitoring," in *Proceedings of the 2005 27th Annual International Conference of the Engineering in Medicine and Biology Society, IEEE-EMBS 2005*, pp. 7161–7164, September 2005.

[18] S. Hamine, E. Gerth-Guyette, D. Faulx, B. B. Green, and A. S. Ginsburg, "Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review," *Journal of Medical Internet Research*, vol. 17, no. 2, p. e52, 2015.

[19] F. Gemperle, C. Kasabach, J. Stivoric, M. Bauer, and R. Martin, "Design for wearability," in *Proceedings of the Second International Symposium on Wearable Computers*, Digest of Papers, pp. 116–122, October 1998.

[20] L. E. Dunne and B. Smyth, "Psychophysical elements of wearability," in *Proceedings of the 25th SIGCHI Conference on Human Factors in Computing Systems 2007, CHI 2007*, pp. 299–302, New York, NY, USA, May 2007.

[21] J. McCann, R. Hurford, and A. Martin, "A design process for the development of innovative smart clothing that addresses end-user needs from technical, functional, aesthetic and cultural view points," in *Proceedings of the 9th IEEE International Symposium on Wearable Computers, ISWC 2005*, pp. 70–77, October 2005.

[22] Attrakdiff, http://attrakdiff.de/index-en.html.

[23] S. Singh and B. Bajorek, "Defining 'elderly' in clinical practice guidelines for pharmacotherapy," *Pharmacy Practice*, vol. 12, no. 4, article no. 489, 2014.

[24] World Health Organization, "Proposed working definition of an older person in Africa for the MDS project," http://www.who.int/healthinfo/survey/ageingdefnolder/en/.

[25] F. J. García-Peñalvo, M. Á. Conde, and V. Matellán-Olivera, "Learning and collaboration technologies. technology-rich environments for learning and collaboration: hci international," in *Mobile Apps for Older Users – The Development of a Mobile Apps Repository for Older People*, vol. 8524, pp. 117–126, Springer International Publishing, Cham, Switzerland, 2014.

[26] L. Qingchuan and L. Yan, "Older adults and digital technology: A study of user perception and usage behavior," in *Proceedings of the AHFE 2016 International Conference on Physical Ergonomics and Human Factors*, vol. 489 of *Advances in Intelligent Systems and Computing*, pp. 155–163, Springer, 2016.

[27] M. Kankaraš, G. Montt, M. Paccagnella, G. Quintini, and W. Thorn, *Skills matter: further results from the survey of adult skills. oecd skills studies*, OECD Publishing, 2016.

[28] M. A. Hentschel, M. L. Haaksma, and T. H. van de Belt, "Wearable technology for the elderly: Underutilized solutions," *European Geriatric Medicine*, vol. 7, no. 5, pp. 399–401, 2016.

[29] A. Rosales, M. Fernández-Ardèvol, F. Comunello, S. Mulargia, and N. Ferran-Ferrer, "Older people and smartwatches, initial experiences," *El Profesional de la Información*, vol. 26, no. 3, p. 457, 2017.

[30] J. E. Lewis and M. B. Neider, "Designing wearable technology for an aging population," *Ergonomics in Design*, vol. 25, no. 3, Article ID 1064804616645488, pp. 4–10, 2017.

[31] J. A. Batsis, J. A. Naslund, L. E. Gill, R. K. Masutani, N. Agarwal, and S. J. Bartels, "Use of a wearable activity device in rural older obese adults: a pilot study," *Gerontology and Geriatric Medicine*, vol. 2, Article ID 2333721416678076, 2016.

[32] K. Mercer, L. Giangregorio, E. Schneider, P. Chilana, M. Li, and K. Grindrod, "Acceptance of commercially available wearable

activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation," *JMIR mHealth and uHealth*, vol. 4, no. 1, 2016.

[33] K. Kunze, N. Henze, and K. Kise, "Wearable computing for older adults - Initial insights into head-mounted display usage," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp 2014*, pp. 83–86, September 2014.

[34] A. Jang, D. MacLean, and J. Heer, "BodyDiagrams: Improving communication of pain symptoms through drawing," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, CHI 2014*, pp. 1153–1162, New York, NY, USA, May 2014.

[35] T. Serif, G. Ghinea, and A. O. Frank, "Visualizing pain data for wheelchair users: a ubiquitous approach," *J. Mob. Multimed*, vol. 1, no. 2, pp. 161–177, June 2005.

[36] E. A. A. Jaatun, D. F. Haugen, Y. Dahl, and A. Kofod-Petersen, "Designing a reliable pain drawing tool: Avoiding interaction flaws by better tailoring to patients' impairments," *Personal and Ubiquitous Computing*, vol. 19, no. 3, article no. A010, pp. 635–648, 2015.

[37] J. Follett, *Designing for Emerging Technologies: UX for Genomics, Robotics, and the Internet of Things*, O'Reilly Media, Inc, 2014.

[38] I. Rodriguez, C. Fuentes, V. Herskovic, and M. Campos, "BePain: A wearable interface to self-report pain and emotions," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp 2016*, pp. 1120–1125, New York, NY, USA, September 2016.

[39] B. A. Rosser, K. E. Vowles, E. Keogh, C. Eccleston, and G. A. Mountain, "Technologically-assisted behaviour change: A systematic review of studies of novel technologies for the management of chronic illness," *Journal of Telemedicine and Telecare*, vol. 15, no. 7, pp. 327–338, 2009.

[40] K. Lyons and H. Profita, "The multiple dispositions of on-body and wearable devices," *IEEE Pervasive Computing*, vol. 13, no. 4, pp. 24–31, 2014.

[41] Y.-M. Fang and C.-C. Chang, "Users' psychological perception and perceived readability of wearable devices for elderly people," *Behaviour & Information Technology*, vol. 35, no. 3, pp. 225–232, 2016.

[42] X. Liu, K. Vega, P. Maes, and J. A. Paradiso, "Wearability Factors for Skin Interfaces," in *Proceedings of the the 7th Augmented Human International Conference 2016*, pp. 1–8, New York, NY, USA, Feburary 2016.

[43] C. Cornelius, R. Peterson, J. Skinner, R. Halter, and D. Kotz, "A wearable system that knows who wears it," in *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys 2014*, pp. 55–67, New York, NY, USA, June 2014.

[44] I. Rodríguez, C. Fuentes, V. Herskovic, and M. Campos, "Monitoring chronic pain: comparing wearable and mobile interfaces," in *Lecture Notes in Computer Science*, vol. 10069, pp. 234–245, 2016.

[45] A. Ferrari, "Digital competence in practice: an analysis of frameworks," Tech. Rep. 9, Research Centre of the European Commission, 2012.

[46] J. Brooke, "Sus-a quick and dirty usability scale," *Usability Evaluation in Industry*, vol. 189, no. 189, pp. 4–7, 1996.

[47] T. Tullis and W. Albert, *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*, Morgan Kaufmann Publishers Inc, San Francisco, CA, USA, 2008.

[48] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[49] B. A. Rosser and C. Eccleston, "Smartphone applications for pain management," *Journal of Telemedicine and Telecare*, vol. 17, no. 6, pp. 308–312, 2011.

[50] G.-N. Gimhae, "Six human factors to acceptability of wearable computers," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 8, no. 3, 2013.

[51] J. Rantakari, V. Inget, A. Colley, and J. Häkkilä, "Charting design preferences on wellness wearables," in *Proceedings of the 7th Augmented Human International Conference, AH 2016*, New York, NY, USA, February 2016.

[52] C. Buenaflor and H.-C. Kim, "Six human factors to acceptability of wearable computers," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 8, no. 3, pp. 103–114, 2013.

[53] J. Van Heek, A. K. Schaar, B. Trevisan, P. Bosowski, and M. Ziefle, "User requirements for wearable smart textiles. Does the usage context matter (medical vs. sports)?" in *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare of Pervasive Healthcare (PervasiveHealth '14)*, pp. 205–209, ICST, Brussels, Belgium, 2014.

[54] A. Pyae, T. Liukkonen, T. Saarenpää, M. Luimula, P. Granholm, and P. Smed, "When japanese elderly people play a finnish physical exercise game: a usability study," *The Journal of Usability Studies*, vol. 11, no. 4, pp. 131–152, 2016.

[55] F. H. A. Razak, R. Sulo, and W. A. W. Adnan, "Elderly user mental model of reminder system," in *Proceedings of the 10th Asia-Pacific Conference on Computer-Human Interaction, APCHI 2012*, pp. 193–200, New York, NY, USA, August 2012.

[56] J. Isleifsdottir and M. Larusdottir, "Measuring the user experience of a task oriented software," in *Proceedings of the International Workshop on Meaningful Measures: Valid Useful User Experience Measurement*, pp. 97–101, 2008.

[57] ISO, "Ergonomics of human system interaction - part 210: human-centred design for interactive systems," Tech. Rep. SO DIS 9241-210, International Organization for Standardization, Genève, Switzerland, 2010.

[58] J. McCarthy and P. Wright, "Technology as experience," *Interactions*, vol. 11, no. 5, pp. 42-43, 2004.

[59] R. D. Ellis and J. C. Allaire, "Modeling computer interest in older adults: The role of age, education, computer knowledge, and computer anxiety," *Human Factors*, vol. 41, no. 3, pp. 345–355, 1999.

[60] S. Kumar, L. C. II. Ureel, H. King, and C. Wallace, "Lessons from our elders: Identifying obstacles to digital literacy through direct engagement," in *Proceedings of the 6th ACM International Conference on PErvasive Technologies Related to Assistive Environments, PETRA 2013*, New York, NY, USA, May 2013.

[61] J. M. Nassar, K. Mishra, K. Lau, A. A. Aguirre-Pablo, and M. M. Hussain, "Recyclable Nonfunctionalized Paper-Based Ultralow-Cost Wearable Health Monitoring System," *Advanced Materials Technologies*, vol. 2, no. 4, Article ID 1600228, 2017.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 7261958, 17 pages
https://doi.org/10.1155/2017/7261958

Research Article

DAFIESKU: A System for Acquiring Mobile Physiological Data

Maider Simón,[1] Ezekiel Sarasua,[1] Borja Gamecho,[2] Edurne Larraza-Mendiluze,[1] and Nestor Garay-Vitoria[1]

[1]*Egokituz Laboratory, Informatika Fakultatea (UPV/EHU), Manuel Lardizabal Pasealekua 1, Donostia, 20018 Gipuzkoa, Spain*
[2]*WimbiTek, Tolosa Hiribidea 72, Donostia, 20018 Gipuzkoa, Spain*

Correspondence should be addressed to Nestor Garay-Vitoria; nestor.garay@ehu.eus

Received 23 June 2017; Revised 7 September 2017; Accepted 28 September 2017; Published 1 November 2017

Academic Editor: Pino Caballero-Gil

Gathering physiological data when they are performing experiments requires a great effort from researchers. Very often, a considerable time is required to prepare the signal acquisition equipment, conduct the experiments, and properly label the data of each participant. Nevertheless this data is valuable for the analysis of personal characteristics, such as behavior, health conditions, and preferences. With the aim of assisting researchers with such tedious tasks, we have developed the DAFIESKU system. This system serves to acquire several types of physiological data. DAFIESKU facilitates the creation of new datasets with physiological data by means of mobile and wearable devices. The usability of the system was evaluated in two case studies in a two-step iterative process. Before conducting the second case study, the whole system was improved using the feedback obtained from the first case study. The results achieved show that usability was enhanced in the second version of DAFIESKU.

1. Introduction

People interact with computing systems in a way that is continuously changing: from batch communication to interactive systems and from interacting at certain times and places with desktop computers to continuous ubiquitous mobile interaction with wearables. New paradigms of human-computer interaction (HCI) have arisen, such as augmented reality, ubiquitous computing, and affective computing, and new applications have been discovered, such as eLearning, eHealth, and eGovernment [1–4].

Initially, specific equipment was used to enable these new applications, but nowadays almost any commercial equipment may serve. The use of mobile smart-phones is common, alone or in combination with other components such as smart-watches, glasses, microphones, and clothes. Most of them have integrated sensors, for example, accelerometers, gyroscopes, temperature sensors, pressure sensors, acoustic sensors, and thermography. With this equipment it is easy to record physiological data from users. These data, if adequately stored and preserved, may be used to train machine learning systems to analyze health conditions, user behavior, characteristics, and other similar topics.

There is a huge number of public online repositories storing datasets [5, 6]. Nevertheless, in some cases researchers need to create new datasets to work with, as the task of acquiring the datasets is an essential step in the research process [7].

The way data are obtained is an important issue directly relating to the validity of the data under consideration. Approaches to recording data are based either in controlled environments such as research laboratories or in uncontrolled environments such as in real life [8–10]. The first approach has the advantage of minimizing noise and recording mainly appropriate data, but its principal disadvantage is that naturalness is lost. That is, subjects may feel uncomfortable in a fairly unfamiliar place for them, probably while being observed by several people, that is, researchers. The latter approach has its advantages and disadvantages contrary to the former: subjects are in their real life contexts and therefore data is generated by more natural reactions, although environmental factors may affect the quality of the recordings. However the popularization of smart-phones and advances with wearable devices have made this option more appealing in the last decade.

In this paper, the authors review which design issues are to be taken into account and present a system named DAFIESKU aimed at acquiring physiological data in a mobile uncontrolled environment. Its main objective is to allow the design of experiments to be applied in real conditions and then to record natural data in real environments instead of recording them in laboratories. This system allows researchers to conduct experiments remotely with the collaboration of participants. The initial versions of the DAFIESKU system and the case studies employed to test them are also detailed here.

This paper is an extension of a paper presented at the UCAmI 2016 conference [11, 12]. The system proposed, aimed at remotely recording physiological data of participants without the help of research assistants, has been developed by the usual iterative development cycle proposed on User-Centered Design [13]. Reference [12] presents, schematically, the first cycle of the iterative development. This new paper extends the information given in the previous one. Moreover, this one presents the second iteration which expands the original framework, focusing on improving the guidance necessary for the user to carry on the experiments with DAFIESKU. New screens and arrangement have been added in the mobile application of DAFIESKU following the recommendations obtained in the first case study involving five participants. A new case study has been performed to test the changes and its appropriateness to meet the goals of the new approach. The results of the aforementioned case study, which involved 17 new participants, are also shown in this paper. New sections and subsections have been included and several figures and tables have been added in order to clarify explanations and to show the results achieved. Two tables outline the lessons learned during both experiments.

2. Related Work

Evaluation and testing of research questions and hypothesis are a recurring subject for human-computer interaction researchers. For this reason, it is not surprising to find an abundance of frameworks, toolkits, and platforms devoted to remote evaluation in the scientific literature [14, 15]. Most often these works are related to the web domain and the usability or accessibility of web interfaces. For instance, RemoTest [16] proposes a platform designed for the remote evaluation of usability and the study of interaction with web applications. RemoTest enables experiments to be set up and data to be collected from participants' web browsers using plug-ins.

Over the last decade, the high popularity of smart-phones and their applications, commonly called apps, have presented an opportunity to carry out remote evaluations in mobile environments. For example, Funf [17] is an extensible framework based in Android which is able to collect and upload data from different sensors embedded in smart-phones with the aim of studying social interactions. There is a version called Funf in a Box [18] that automatically creates a phone-sensing app with the guidance of a user. Aware [19] is a framework devoted to the instrumentation of the smart-phone to understand human behavior better. The framework is oriented to context-awareness and it provides mechanisms to interact with the user while acquiring data from the sensors.

In recent years, the use of physiological signals has emerged in the computer science area due to the vast amount of wireless wearable devices that provide such signals. Hence, similar systems for the recording and storing of physiological signals can be found in the literature. Biosignal Ignitor toolkit [20] introduces a set of tools for the acquisition of different physiological signals. The toolkit provides a desktop application to visualize and record the signals and also includes its own sensor platform called BITalino [21]. Focused in the medical domain, Physiodroid [22] presents a framework for the creation of medical applications using sensors including wearable devices. This framework is oriented to developers who want to create their own application for the medical or context-aware domain rather than researchers who want to conduct experiments. These works are focused on the remote acquisition of physiological signals without taking into consideration the whole experimentation process.

There are other paradigms such as *Participatory Sensing* which "will task deployed mobile devices to form interactive, participatory sensor networks that enable public and professional users to gather, analyze and share local knowledge" [23] or *Mobile Crowd Sensing* which is "a new sensing paradigm that empowers ordinary citizens to contribute data sensed or generated from their mobile devices, aggregates and fuses the data in the cloud for crowd intelligence extraction and people-centric service delivery" [24]. These paradigms are closely related to this proposal and could be easily merged with it.

As an example of these paradigms, [25] proposes a system where crowdsourced physiological data and subjective emotions could be used for cross-validation of the emotions that people feel when visiting different urban spaces. However this mobile app is being considered for compiling subjective emotions not for collecting physiological data. This is precisely where DAFIESKU could be of help.

Another crowd-sensing platform is VITA [26], a mobile cyber-physical system for crowd-sensing applications. Similarly to DAFIESKU, VITA software architecture is divided into two parts, one for the mobile part and the other for the cloud. VITA allows human resources to be allocated and supports user participation in various mobile crowd-sensing applications through interaction with other participants. On the other hand, users with a participant role in DAFIESKU interact only with the mobile application following the tasks designed by the researcher.

The main objective of the DAFIESKU system is to provide resources and tools to researchers in order to facilitate the deployment of an experiment from its design to the analysis of the results. Experiments will be made in real or natural environments (familiar to participants) by using wearable technology and ubiquitous computing. The data obtained is expected to be more naturalistic than that obtained in controlled places, such as research laboratories (unfamiliar to participants).

Therefore, DAFIESKU can be used in research studies to acquire physiological data, including studies in medical settings in order to monitor personal health conditions and

in sports settings or even when studying at home in order to measure effort and enhance efficiency. It can also be used for affective computing when detecting emotions [27].

3. Design Issues

There are different types of experiments involving physiological data in computer science research. DAFIESKU aims to collect physiological data with activity labels from the participants of the experiments. Furthermore, the system must work autonomously without the presence of research assistants when the experiments are conducted. Subsequently, the system was designed to interact directly with the experiment participants. In this section design issues on the development of the system (see Figure 1) are explained in deeper detail.

3.1. Involved Roles. When developing DAFIESKU, two main user roles have been identified. The first role is the researcher or the designer of the experiments, while the second one is the experimental participant.

The role of the researcher is to think up, design, and supervise experiments in which real users are going to be involved. They usually work in a team, but they may also work alone. The researcher comes up with research questions and the way to measure responses to them. They also decide the apparatus to be used in the experiment and estimate the duration of the experimentation and the size of the population involved. The methods to obtain data (e.g., videotaping, interviewing, recording speech), where to store them, and whether to make them available or not for the community are also decisions to be taken by the researcher. It should be underlined that DAFIESKU is aimed to help in the setting up of experiments, the selection of participants, and the storing of data. Analyzing data is beyond the scope of DAFIESKU.

On the other hand, there are the experiment participants. They have to decide whether they are willing to collaborate with researchers in order to create data that will be further analyzed. They have to explicitly give their consent and, if requested, they and their data are to be deleted from the experimentation. Experiments are made individually or in groups, depending on the design made by the researchers. Participants are also known as experimental subjects.

3.2. Experimental Design. Researchers usually take several steps in order to define experiments and the experimental setups to be carried out with final users [9]. Experimental research analyzes and explains how manipulating a known variable ("independent variable") has a direct influence on another variable ("dependent variable"). There also are other types of variables named "control variables" which are to be kept constant across the whole study in order to ensure that the achieved results are indeed caused by the independent variable.

The main types of experiments are those called "randomized experiments" and "quasi-experiments." In randomized experiments, participants are randomly assigned to every independent variable condition while in quasi-experiments participants are manually assigned to experimental conditions.

Randomized experiments are typically between-subjects or within-subjects. In between-subject experiments, participants are assigned to a single group which is then exposed to a single condition. If the number of conditions is high, the number of required groups increases and so in turn does the number of required participants. In within-subject experiments participants are exposed to every experimental condition and it is easier to find differences caused by the treatment. The number of required participants is lower than in between-subjects experiments, but a learning effect could appear. In order to minimize these effects, conditions are usually counterbalanced (complete counterbalance and/or Latin square).

Another type of experiment is the Factorial Design, used when more than one independent variable has to be tested. Factorial experiments can be within-subject, between-subject, and mixed (the last one, for example, when testing treatments related to age and/or gender).

In DAFIESKU, we have been focusing on experiments to get data from participants, especially physiological data. Therefore, we aim to provide researchers with an application that will help them while designing their experiments to obtain physiological data.

3.3. Making Experiments Available to Participants. When experiments are to be made in controlled environments such as research laboratories in presence of researchers, these can give all the information to participants directly. However, when researchers and participants are in different locations, a protocol of collaboration needs to be established.

In DAFIESKU, there are these uncontrolled environments which form the principal focus, for example, the participant's habitual dwelling. In this environment, participants feel more comfortable, responding as they might do in their day to day lives, and the recorded data is more realistic, reflecting user's natural characteristics and behavior.

The experiment, its related information, and the required equipment (e.g., sensors) have to be made available for participants. This is achieved by means of their mobile phones, and the data is sent through the same phones. Therefore, when a participant is ready to perform an experiment, first they have to open the application which contains the experiment, next follow the steps detailed in the experiment, and, finally, send the data captured using sensors.

3.4. Wireless Sensors, Data Acquisition, and Transmission. Nowadays, sensors are more comfortable due to miniaturization. Normally, these sensors are integrated in all-in-one hardware platforms, such as Biosignalsplux [28] or Shimmer [29], including batteries and wireless transmission capabilities. In short, these sensor platforms are useful for experiments performed in uncontrolled environments, for instance, at the home of the participant.

Nevertheless, these platforms still have limited storage capabilities and they use Personal Area Network (PAN) connectivity, such as Bluetooth, which can not be connected directly to the Internet.

In order to overcome these limitations, a common solution is to combine sensor platforms with mobile phones

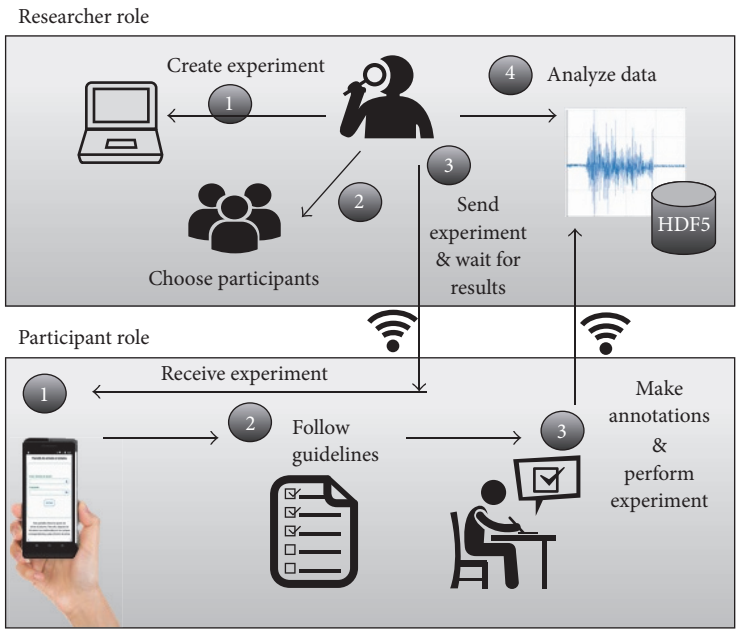


Figure 1: Design issues addressed with DAFIESKU.

carried by users. Thanks to the mobile devices, the data obtained from the sensors are transmitted first to the storage of the mobile device and finally to the system server which may be available on the cloud.

3.5. Data Storage and Analysis. Once all the ethical and privacy issues have been taken into account [30], recorded data should be processed and filtered and, in order to make further analysis easier, standards, such as HDF5, [31] should be followed.

The physiological data must be available in the system, not only for the researcher but also for the participant. Access to their own physiological data contributes to ensuring transparency in the process of data acquisition.

Of course, if the data is to be used by different research teams, various agreements have to be made between the owners or administrators of the databases, in order to ensure that researchers do not make illicit use of these data when processing them. For instance, the anonymity of users involved in the recording of data has to be preserved even when the data are made publicly available.

3.6. Technical Details. The server side of DAFIESKU was developed in a VirtualBox virtual machine with the Ubuntu 14.04 operative system. Various technologies are used to implement the different parts of the server side of the system. The server configuration is carried out with Apache 2.4.7. The web application is implemented with HTML5, PHP 5.5.9, JavaScript (jQuery 3.0.0), and CSS3. To implement the REST API, the Slim micro framework 2.6.3 is used, which provides an easy way for the Android device to access the data stored in the database. The protocol used to carry out the various requests is HTTP. The JSON format [32] is used to facilitate the communication between the server and the Android smart-phone/tablet by means of the API.

Depending on the request, different JSON objects have to be sent to the API. For example, when a participant finishes an experiment, the Android APP will send a request to the server via the API to update the information on that specific experiment in the database. Examples of the parameters needed for that given operation are the id of the experiment and the data that needs to be updated. If all the required parameters are sent by the Android device when calling the API, the operation is successfully carried out and a specific JSON response is sent to the Android device saying that it has been completed. The JSON response from the server is always formed by an HTTP status code and a message; this could be an error message, a message saying that the operation was successful, and/or a nested JSON message containing the data requested.

MySQL 5.5.47 is used for database management. All the experiment data are stored in the database, so it can be accessed by the API. We also needed a way to store the physiological data, so we opted for the HDF5 file format. When a participant finishes an experiment, the Android device automatically creates a compressed ZIP file containing the physiological data obtained (one text file per sensor used and another file for the time marks) and uploads it to the server using a specific request method developed in the REST API. The server then runs a program developed in C that uses the HDF5 1.8.16 library, which reads the ZIP file containing all the data, and converts it to an HDF5 format file. This file is then stored in the server and its directory path is linked in the database to the experiment to which it belongs, so the researcher has the file available for download.

The application for the Android client of DAFIESKU has been developed using the Android Studio IDE [33] and it can

be run on Android devices up to API 14 (Android 4.0, Ice Cream Sandwich).

Most of the classes used for the project are Java and Android native, although some third-party libraries have been used, such as Butter Knife [34] for field and method binding, and GraphView [35] for plotting in Android views. Both libraries are open source and widely known by Android developers.

As mentioned, the communication between the Android smart-phone/tablet and the physiological data acquisition device will be performed through the Bluetooth protocol. Therefore, the Android library provided by the physiological data acquisition device should be added to the application project. In this particular scenario, a BITalino library for Android has been added. If the physiological data acquisition device does not have an Android compatible library the communication with the Android device can be manually performed using the Android Bluetooth class [36].

On the other hand, the Android device and server communication is performed using the HttpURLConnection [37] Java class following the JSON [32] format. Four different types of calls are used by the Android application to communicate with the server:

(i) The login call: in this call credentials introduced by the user are sent to the server. If login succeeds an API key will be received as a response. This API key will be used to verify the user in further communications.

(ii) Synchronization call: this call will update the Android application with the latest information relating to tasks saved in the server; for example, it will download to the app new tasks available to the user.

(iii) Sending the extracted data: this HTTP POST call will be executed automatically when a task is successfully finished. It sends the data extracted from the physiological sensors to the server.

(iv) Confirming call: once the data transmission (previous call) is finished an HTTP PUT call will be performed to update the task as carried out on the server side.

In future versions, for these communication purposes, it is planned to add the Retrofit library [38]. This will make the application easier to maintain and will increase stability.

The data acquired from the physiological sensors will be stored in temporary files inside the device's internal memory, one file per sensor. Due to the computing limitations of mobile devices, generated files will not be processed in the client. Instead, all the temporary files will be compressed in ZIP format file and sent to the server to be further processed. Once transmission with the server is confirmed the data will be removed from the Android device.

4. DAFIESKU 1.0 System

The DAFIESKU system has been designed to make use of wireless devices to obtain a less intrusive environment for the participants during data acquisition. DAFIESKU consists of a set of wearable devices with physiological sensors coordinated by a server. Wearable devices with wireless connections usually offer Bluetooth based connectivity which can be used to send physiological data to any computer. For this reason, a mobile device (e.g., smart-phone or tablet) is used as an intermediary between the server and the wearable devices (see Figure 2).

The main advantage of this approach is the possibility it affords of configuring the data acquisition in each experiment and the possibility of guiding and supporting the participants during the experiments. It should be pointed out that both researchers and participants must register on the DAFIESKU system. Usually, the participants will be registered on DAFIESKU by the researchers.

Prior to acquiring or recording data, there are several decisions to be adopted during the configuration phase. Using a web interface the researcher prepares the experiment which the participants will be able to download to their mobile devices, thus enabling the researcher to acquire their physiological data. In this section, we present the first version of DAFIESKU, while in Section 6 we present the new version that has been developed considering the results obtained in the experimentation presented in Section 5.

4.1. Creating Experiments. With DAFIESKU, researchers may establish which data will be acquired and how and for what length of time it will be acquired in order to create databases with physiological information.

The data to be gathered depends on the scope of the studies that the researchers plan to develop in the future, the sensors that are available for acquisition, the population, and so forth.

The DAFIESKU system provides the researchers with a web client, which enables them to define how the sensors (among those available) will be used, whether they will be used together or separately, and which kind of information provided by the devices will be recorded (see Figure 3).

Figure 3 shows a form for creating an experiment in the DAFIESKU 1.0 web application. The values requested are a name for the experiment, not necessarily unique (there can be more than one experiment with the same name, as they are identified by an id number), and a description, which should briefly and clearly explain its purpose to the participant. Next, the sensors that are going to be used to get the physiological data must be selected. Afterwards, the duration of the experiment (in hours, minutes, and seconds) and the name and surname of the associated participants have to be defined.

Finally, several time marks can be defined. These time marks represent extra information such as activities that a participant can perform during an experiment, temporary emotions, and so forth. For testing reasons, in this version of DAFIESKU 1.0, a maximum of three time marks can be defined.

4.2. Providing Guidance to the Participants. In the DAFIESKU system, participants play a key role in data acquisition. Among other things, they decide when to start the experiments and they provide feedback with which to label the physiological data.

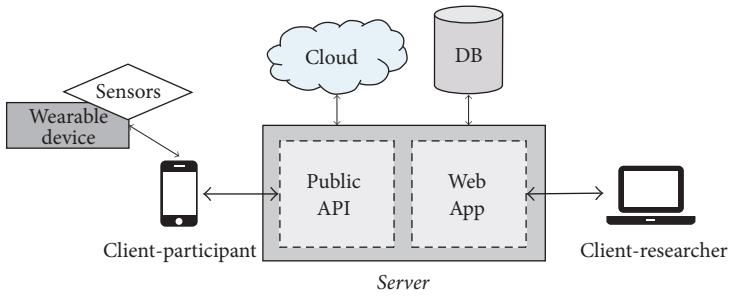


Figure 2: Architecture of the DAFIESKU system.

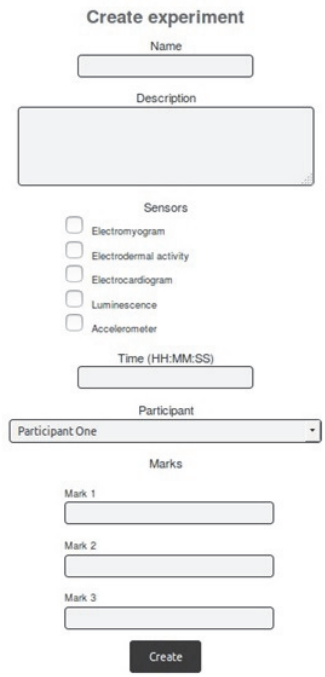


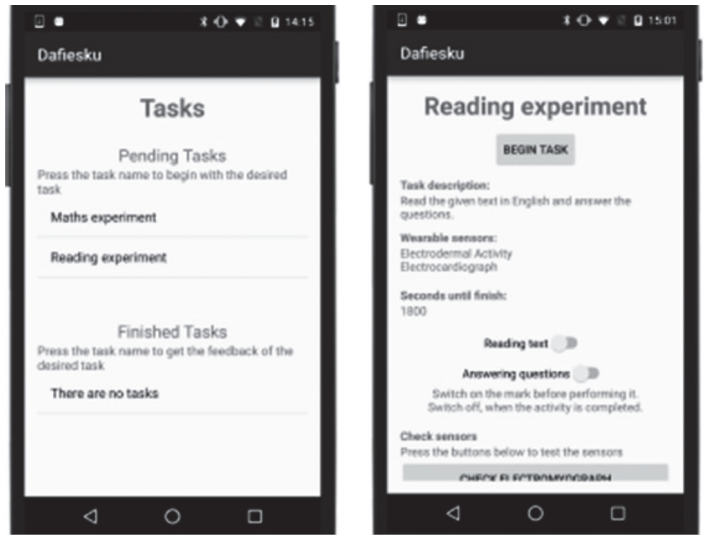
Figure 3: Designing an experiment with the web client of DAFIESKU 1.0 system.

In order to facilitate these activities, the researcher must also (1) specify the experiment description and define the tasks for the participants, (2) describe the instructions for attaching the sensors, and (3) set a list of predefined marks or activities related to the experiment to label the resultant data set. All these texts must be comprehensive and understandable for the participants in order to avoid problems during experimentation (see Figure 4).

4.3. Data Acquisition during the Experiments. The participants will be able to download the experiment to their Android mobile device (smart-phone or tablet), thanks to a communication that uses a web service based on a REST API. Once downloaded, they will have to set the sensors and connect them wirelessly to the mobile device, that is, via Bluetooth. Finally, they will have to start to record data and to label moments, following the instructions of each experiment. These last two will be mere *push button* or *select from list* tasks, respectively. Therefore, in the first case untagged data are recorded when the user wants to, whereas in the second case data are recorded with marks relating to the activity the user is carrying out (such as *running*, *climbing slopes*, *studying*, and *being relaxed*).

Figure 4(a) shows the main activity of the DAFIESKU Android application. The aim of this screen is to show all the tasks relating to a previously logged user. Tasks are classified as being pending or finished, depending on their status. In order to perform a task, it has to be selected from the pending tasks list. By clicking on a finished task name the participant can obtain the task feedback.

Once the user has selected a task from the pending list, the screen in Figure 4(b) is shown. This screen contains all the information regarding the selected task (description, sensors to wear, maximum time estimation, and activities to perform). In order to make correct recordings, the user can check whether sensors are correctly adjusted before performing activities related to the task. When participants complete the tasks before the finish time estimation, the



(a) (b)

Figure 4: Mobile application screenshots. (a) shows pending tasks and (b) shows the experiment description.

back button will take the participant to the previous screen. Otherwise, when the estimated time is reached, the system will automatically go back to the previous screen.

5. Case Study to Evaluate the DAFIESKU 1.0 System on Non-Classroom Learning Experimentation

A case study was carried out to test the usability of the DAFIESKU 1.0 system. For this purpose a representative experiment in a non-classroom learning situation was designed for DAFIESKU in order to test both roles in the system: that of the researcher and that of the participant. As a researcher, the experimental subject introduces the information to set up the experiment. On the other hand, as an experiment participant, the same experimental subject performed fixed tasks such as reading texts or questions, and answering questions while their physiological data were acquired. The usability of both web and mobile applications was measured by using SUS questionnaires [39].

5.1. Method

5.1.1. Participants. 5 volunteers (2 females) were recruited from the surrounding research laboratories of the Faculty of Informatics of the University of the Basque Country (UPV/EHU). This is the minimum number of participants that lets you find almost as many usability problems as you would find using many more test participants, as indicated in [40]. The experimental subjects ranged from the age of 30 to 57 years (41.2 ± 10); see Table 1. Informed consent was obtained from all individual experimental subjects included in the study. Users were computing graduates except User 05a, who had a degree in industrial engineering. Users 01a, 02a, and 05a had PhDs. All of them had previously participated in other research studies, but only Users 03a and 04a had been previously involved in physiological experimentation. Furthermore, Users 03a and 04a had previously prepared experiments with users, but no physiological data had been captured on those occasions.

Table 1: Participants related information.

User	Age	Gender
User 01a	50	Female
User 02a	30	Male
User 03a	36	Male
User 04a	33	Male
User 05a	57	Female

5.1.2. Apparatus. The web service for designing experiments runs on a virtual machine with an Ubuntu 14.04 LTS operating system. This web service was used when playing the researcher role in this experiment using a laptop. The designed apps for acquiring data run on a Samsung Galaxy Tab 2.7.0 with an Android 4.2.2 operating system (RAM: 1 GB; 1.0 GHz dual core processor; 7 screen with $1024 * 600$ pixels). The BITalino [41] sensor platform was used to collect physiological data from the electrocardiography (ECG) and electrodermal activity (EDA) sensors. The virtual machine running on the laptop and the tablet were connected wirelessly using a WiFi router. The tablet and the BITalino were connected via Bluetooth interface.

5.1.3. Procedure and Design. After consent was obtained from the experimental subjects, their demographic data were gathered. Then, a sheet with the tasks to be completed was delivered to each experimental subject.

Table 2: Time (in minutes) needed to complete tasks.

User	Task 1	Task 2	Task 3	Total
User 01a	5′	21′	1′	27′
User 02a	7′	34′	3′	44′
User 03a	4′	21′	2′	27′
User 04a	5′	26′	2′	33′
User 05a	8′	16′	1′	25′
Mean	5.80′	23.60′	1.80′	31.20′
STD	1.47	6.09	0.75	6.94

Table 3: Scores achieved in SUS questionnaires.

User	Researcher	Participant
User 01a	77.5	80
User 02a	65	77.5
User 03a	67.5	42.5
User 04a	92.5	87.5
User 05a	90	67.5
Mean	78.5	71
STD	11.25	15.62

The first task (Task 1) to be completed was to define an experiment with certain characteristics by using the web client of DAFIESKU 1.0. This was carried out in the researcher role. Task 2 was carried out in the role of a participant. The equipment needed in the experiment was prepared (see Figure 5), it was verified that everything was working correctly, and then the activities required by the experiment defined in Task 1 were carried out. This experiment consisted of reading a text printed on paper and then reading and answering several questions relating to that text, also on paper and using a pen, as if they were doing out of classroom exercises. Participants had to indicate by using the system the time in which they were reading the text, when they were reading questions, or when they were answering questions. After the experiment was finished, they had to take off the sensors. As a researcher again, on Task 3 they verified, using the web client again, that the data of the experiment was properly stored on the server.

When finishing these three tasks, experimental subjects completed two SUS questionnaires, one as researcher and the other as participant in the physiological experiment. The time required to complete each of the tasks was measured. Finally, they were interviewed by up to two DAFIESKU 1.0 development team members in order to get more feedback, in this case qualitative feedback. The interviews were around 5 minutes long and the questions referred, in particular, to their opinion, their suggestions, the usability of the system, and any issues they had had during the process, as well as what aspects might be improved for future versions. Moreover, the interviews were voice recorded.

5.2. Results and Discussion. All the experimental subjects successfully completed the required tasks. Table 2 shows the times needed to complete the three tasks. Task 2 required the most time while Task 3 was the one requiring the least time. Minutes needed for completing the three tasks range from 25 (User 05a) to 44 (User 02a).

Concerning the usability of the system (see Table 3), the SUS scores were 78.5 ± 11.25 for the researchers web client application and 71 ± 15.62 for the Android application. These scores are over 70 and therefore both applications may be considered as being good from the usability point of view [39]. Nevertheless, two of the participants marked the applications lower than 70 and one of them rated the Android application with a score of 42.5, indicating that there were usability concerns to be addressed.

Experimental subjects also suggested several enhancements, almost all with the aim of improving participant experience with regard to data acquisition. One of them (User 03a) suggested including instructions about how to wear the physiological sensors required for the experiment app. Some other experimental subjects mention that wires hindered the completion of the experiment during Task 2. As DAFIESKU aims to create experiments and register data, including help or instructions for the correct adjustment of sensors directly in the app was considered a good idea for future versions. However, it has to be said that coping with wires depends on the sensor technology employed and it is beyond the scope of the DAFIESKU system. Nevertheless, the researcher should consider the characteristics of the technology employed when designing experiments for preventing usability issues that may influence the results.

Other comments advocated the enhancing of the DAFIESKU 1.0 implementation. For example, with relation to the mobile application, several participants suggested increasing the distance between task names when they are shown on the screen. They also suggested maintaining the vertical orientation of the screen while they were checking whether sensors were correctly working. With relation to the server-side application, how to express the duration of the experiments and how to select sensors involved in the experiment needed to be more intuitive. And in order to minimize errors when designing experiments, a confirmation option was suggested to allow the researchers to read the details and accept the experiments after checking if they were correct.

In Table 4 lessons learned in order to make new versions of the systems are highlighted, relating to both the server-side application and the mobile application.

There were also several issues which brought to light aspects in which further evaluations of DAFIESKU 1.0 could be enhanced. For example, our case study was made with five participants all of whom had technological profiles. For more complete and significant results, DAFIESKU should be evaluated with more participants with different profiles (for example, psychologists or people with medical profiles).

6. DAFIESKU 2.0 System

In this section the main changes with respect to the first version of the DAFIESKU system are explained, taking into

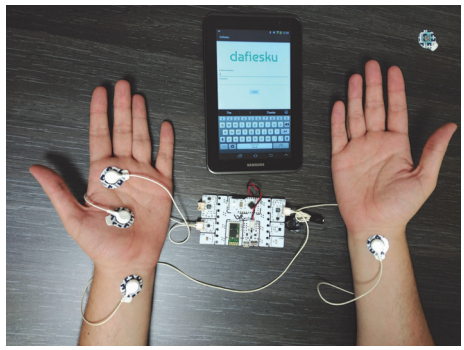


Figure 5: Experimental setup: the role of participant.

Table 4: Lessons learned after case study 1.

#	Server-side application
(1)	Make intuitive the expression of the duration of experiments.
(2)	Make intuitive the selection of the involved sensors.
(3)	Include a confirmation option to check whether the experiment has been correctly defined.
(4)	Maintain the ease of use and the low complexity.
(5)	In general, it has been found that most people would design experiments quickly with this system.
#	Mobile application
(1)	Include instructions about how to wear physiological sensors.
(2)	Increase the distance between task names on the screen.
(3)	Ensure low complexity of the app to enhance the ease of use.

account the results extracted from the previous experimentation.

6.1. Creating Experiments. With DAFIESKU 2.0, some improvements have been made to the web client in order to solve certain usability issues, also following the suggestions made by the participants that took part in the previous experimentation (see Section 5).

In this 2.0 version of DAFIESKU web client, the objective is maintained, which is to facilitate the process of creating experiments, but in a more intuitive way.

Figure 6 shows a form with which an experiment can be set up in the DAFIESKU 2.0 web application. The main differences are the order of the parameters required to create an experiment. As the participants need to be registered in the system by the researchers, before specifying other parameters for the experiment, it is important to check whether the participant is registered in the system or not. For this reason, in DAFIESKU 2.0, the first step is that of selecting the participants.

Furthermore, more sensors have been added as well as the possibility of selecting a different device. As the system is generic, DAFIESKU 2.0 is prepared to work with different devices. This also facilitates the connection and use of additional devices that allow the use of different sensors for the experiments. When creating an experiment, the system will indicate to the researcher, in real time, which devices are compatible with the sensors that they have selected.

Next, the maximum duration of the experiment must be defined, in hours and minutes. This field has been also changed from the 1.0 version. In fact, the participants had issues with the time format they needed to specify, which lead to errors when creating the experiments. This field has been simplified in order to avoid this kind of error. Moreover, the field for specifying the seconds has been removed, as it makes little sense outside of testing purposes.

Finally, certain time marks can be defined which represent the different activities a participant can perform during an experiment. In DAFIESKU 2.0, it is possible to define dynamically as many time marks as needed by the researcher, unlike in version 1.0, in which only up to three time marks could be specified.

6.2. Improved Guidance for the Participants. As mentioned in previous sections, the aim of the mobile application consists in the recording of physiological data. Due to the critical nature of these data, it is advisable to give the user some feedback to make sure that they are performing the recording correctly. This is even more important if the user has no experience of mobile applications as is the case of several elderly people. For all these reasons and taking into account the case study of the previous section, in the new version of the mobile application, four different categories of assistive screens have been included to guide the user (see Figure 7):

(1) Default application information: this category describes the general information about the mobile application such as how to log into the system and what a user has to do to start an experiment (see Figure 7(a) for login screen guide). These screens are

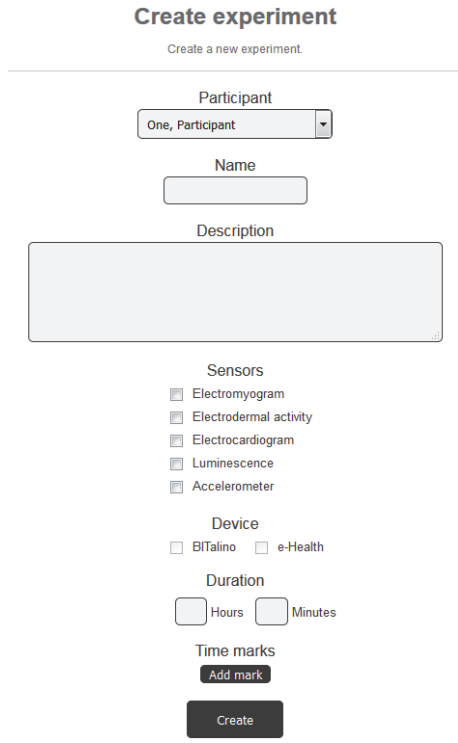


Figure 6: Designing an experiment with the web client of DAFIESKU 2.0 system.

accessed automatically the first time a participant opens the application and after that moment by means of a small information icon at the top of the screen (see Figure 8(a)). This button is implemented for all the screens providing in each one the appropriate information to be able to continue using the application.

(2) Device functionality and sensor placement assessment: for an inexperienced user, physiological data acquisition platforms may be complicated to use for the first time. Furthermore, the placement of electrodes or the platform itself in the body requires a careful explanation to avoid errors when recording the data. A step-by-step assessment is provided for the device and all the electrodes required in the experiment (see Figure 7(b) along with a guide in Figure 7(c) for both the device and electrode placement).

(3) Sensor testing screens: once the sensors are attached to the participants body, a series of tests should be performed to determine that the sensors are working as expected. The system achieves this goal showing to the participant two figures: first, the expected wavelet for a physiological signal and, then, their own signal recorded by the application in real time (see Figures 7(d) and 7(e) for electrocardiogram signal test). The system asks the participant to continue or fix the placement of the sensors if any error was found.

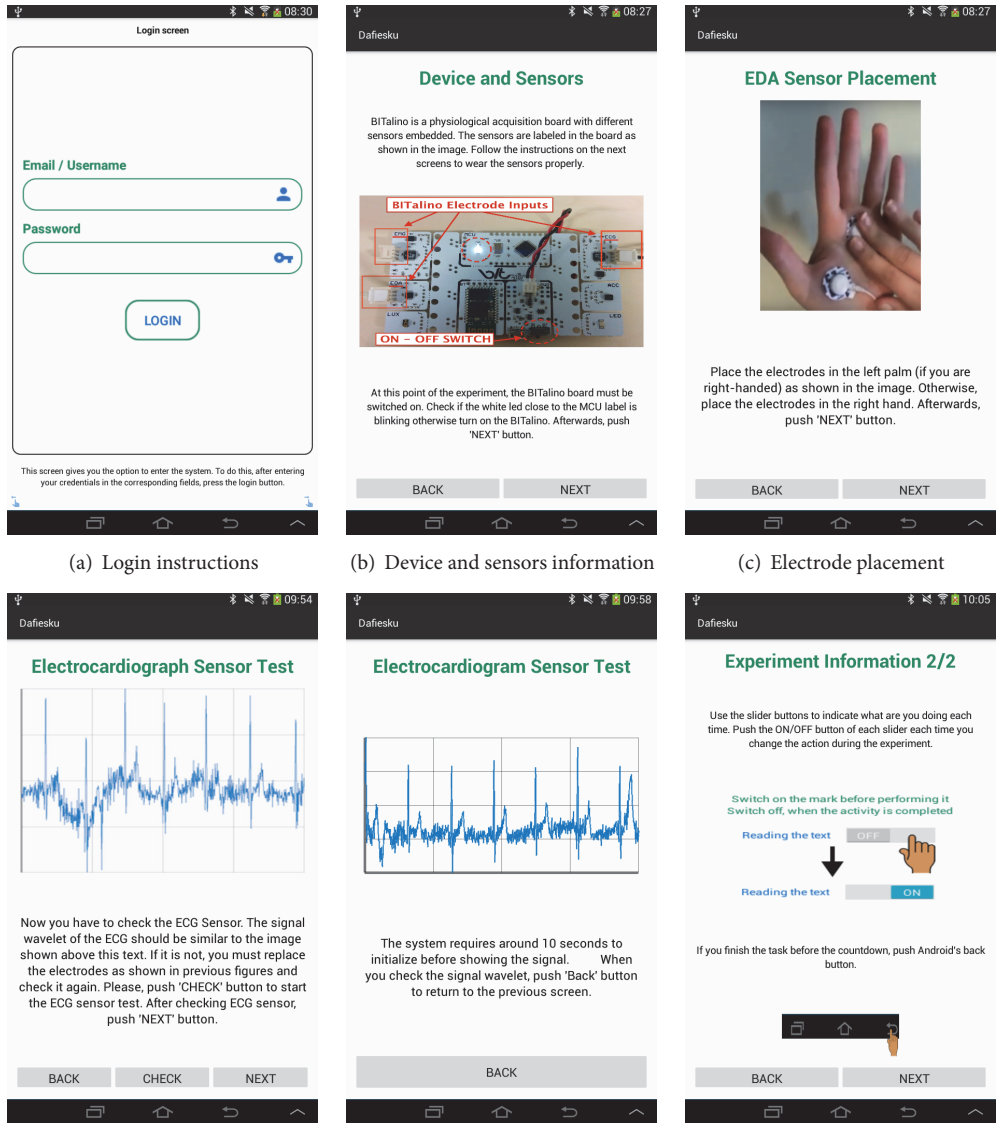
(4) Experiment screen instructions: the system was developed to annotate various actions during the experiments using predefined markers. A brief explanation with the functionality of the annotations is provided to the participants (see Figure 7(f)).

The first category of assistance screens is accessible at any time via an icon. On the other hand, the rest of the categories are accessed sequentially when an experiment starts. This sequence must be carefully followed by the participants to prevent some common mistakes when using physiological sensors. Overall, more than 15 different screens were introduced for the 2.0 version of the mobile application.

6.3. Usability Improvements in the Mobile Application. After the first case study, we noticed that the usability of the mobile application could be enhanced. In DAFIESKU 2.0 several improvements relating to the user experience were added.

Thus, fields that do not require interaction with the user are now colored differently on the screen to those that do require interaction. Moreover, for DAFIESKU 2.0 several images were added in the fields in which the user has to interact in order to give a more graphic description, for example, an image of a key in the password field (see Figure 8(a)).

Furthermore, another significant change was added: an update button on the main screen (see Figure 8(b)). By clicking on this button the user synchronizes the status of his local tasks with those on the server. In DAFIESKU 1.0, this



(a) Login instructions (b) Device and sensors information (c) Electrode placement

(d) Sensor signal instructions (e) Sensor test with real signal from participant (f) Experiment screen instructions

Figure 7: Mobile application guidance screens.

synchronization was only carried out at the time of setting up or finishing an experiment.

Finally, several internal improvements were implemented for the mobile application, taking into account the first experiment and the way users interacted with the mobile application. Therefore, the implementation was slightly adapted to give greater stability to issues which were noticed during the experimental phase. For example, a more efficient keyboard management was included.

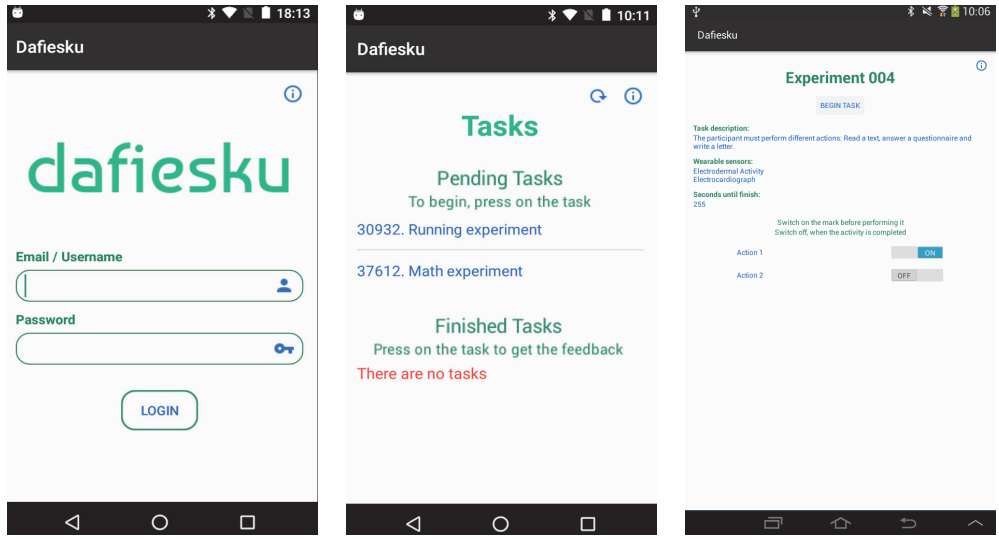
7. Second Case Study: Evaluation of DAFIESKU 2.0

The newly developed version of the system includes enhancements on the server side and in the mobile application that must be evaluated again to measure if any improvement was achieved in terms of usability (i.e., personal satisfaction and efficiency). We address two different research questions in this second case study:

(i) Q1: how much more usable is the second version of the web application than the previous version of the web application?

(ii) Q2: how much more usable is the second version of the mobile application than the previous version of the mobile application?

The same protocol followed in the previous case study was replicated for the second case study. Therefore, the apparatus, procedure, and design subsections of the Method section



(a) Login (b) Experiment list (c) Experiment running

Figure 8: Mobile application experiment screens.

remain the same as the ones described in Section 5 and they should be considered as control variables.

In order to answer research questions Q1 and Q2, the completion times for each task and the SUS scores for each role were collected. In addition to these measures, the participants were requested to fill in a Likert-scale questionnaire with values from 1 to 5 points at the end of the experiment. The questionnaire contains 8 items which are described in Table 5.

Finally, they were interviewed by a DAFIESKU 2.0 development team member in order to get more feedback. The interviews were around 10–15 minutes long and the questions made were especially regarding their opinion, suggestions, usability of the system, and issues they had during the process, as well as what aspects should be improved for future versions. The interviews were voice recorded.

7.1. Method

7.1.1. Participants. 17 volunteers (7 females) were recruited from the Faculty of Informatics of the University of the Basque Country (UPV/EHU) and neighboring faculties. The experimental subjects ranged from the age of 21 to 51 years (30.65 ± 8.8); see Table 6. Informed consent was obtained from all the individual experimental subjects included in the study. In general, users had a degree in computing, except User 06b who had a degree in mathematics, Users 07b and 17b with degrees in electronics and User 12b, who was studying psychology. Moreover, Users 01b, 09b, 15b, and 16b were Ph.D. graduates. Users 02b and 05b were left-handed while the others were right-handed. Users 01b, 03b, 09b, and 16b recounted that they had prepared experiment(s) with users in the past, and User 09b had previously prepared at least one experiment in which physiological information was acquired. Several users had had previous experience as participants in experiments (Users 01b, 03b, 05b, 06b, 09b, 10b, 12b, 13b, 14b, 15b, 16b, and 17b), and in some cases their physiological data had been acquired in at least one previous experiment (Users 06b, 09b, 12b, 13b, 14b, and 17b).

7.2. Results and Discussion. All the experimental subjects successfully completed the required tasks. Table 7 shows the times needed to complete the three tasks presented. Task 2 was again the one needing more time while Task 3 was the one which required the least time. The number of minutes needed to complete the three tasks ranged from 17 (User 04b) to 38 (User 11b).

Comparing these results with the previous evaluation (Section 5), there have been improvements in Task 1 (4.06′ on average while the previous study achieved 5.80′ on average), Task 2 (19.65′ on average while in the previous study 23.60′ was needed on average), and in Total (26.06′ on average compared to 31.20′ on average in the previous study), while Task 3 achieved poorer results (2.35′ on average while in the previous study an average of 1.80′ was achieved).

With relation to the usability of the system (see Table 8), the SUS scores were 81.18 ± 9.81 for the researchers web client application and 80 ± 13.64 for the Android application. These scores may be considered quite good, nearly excellent from the usability point of view [39]. Nevertheless, there are usability concerns to be addressed after considering the individual scores.

When compared to the results shown in Section 5, both components of DAFIESKU have achieved better scores. First, the researchers web client application scored 78.5 on average in the previous study, whereas the Android application achieved 71 on average. The improvement was notable in the case of the Android application.

Moreover, the time required to complete the tasks was lower and the usability results were higher. These data suggest

Table 5: Final questionnaire to be answered by participants.

#	Factor	Statement
(1)	Satisfaction	I feel satisfied because the experiment has been easy for me
(2)	Comfort-1	I felt comfortable using the mobile app anytime
(3)	Autonomy	I am able to make experiments by using the DAFIESKU system without being aided by technical staff
(4)	Comfort-2	Sensors do not cause inconvenience when doing the experiment
(5)	Engagement	I enjoyed it and I think a system such as DAFIESKU is useful to recording physiological data easier
(6)	Instructions usefulness-1	The instructions regarding how to wear the sensors were useful
(7)	Instructions usefulness-2	The instructions for testing and checking sensors were useful
(8)	Instructions usefulness-3	The instructions regarding how the application works were useful

Table 6: Second validation participants related information.

User	Age	Gender
User 01b	51	Female
User 02b	25	Male
User 03b	34	Female
User 04b	26	Male
User 05b	27	Female
User 06b	23	Male
User 07b	25	Male
User 08b	24	Male
User 09b	39	Male
User 10b	28	Female
User 11b	30	Male
User 12b	21	Male
User 13b	30	Male
User 14b	29	Female
User 15b	32	Male
User 16b	51	Female
User 17b	26	Female

Table 7: Time (in minutes) needed to complete tasks in the second experiment.

User	Task 1	Task 2	Task 3	Total
User 01b	7′	26′	4′	37′
User 02b	3′	18′	3′	24′
User 03b	7′	24′	2′	33′
User 04b	2′	13′	2′	17′
User 05b	3′	18′	4′	25′
User 06b	3′	21′	1′	25′
User 07b	4′	20′	3′	27′
User 08b	3′	19′	2′	24′
User 09b	3′	18′	2′	23′
User 10b	5′	23′	3′	31′
User 11b	4′	32′	2′	38′
User 12b	4′	17′	2′	23′
User 13b	5′	18′	2′	25′
User 14b	2′	17′	2′	21′
User 15b	4′	20′	2′	26′
User 16b	4′	14′	2′	20′
User 17b	6′	16′	2′	24′
Mean	4.06′	19.65′	2.35′	26.06′
STD	1.52	4.61	0.79	5.67

that DAFIESKU 2.0 is more usable, more efficient, and friendlier. Therefore we can conclude that the changes introduced have had a positive effect on the system. With respect to the first research question, the SUS score slightly increased in the web application which indicates that the changes proposed had a minor effect in the usability perception. Additionally, with respect to the second research question, the novelties introduced in the mobile application increased the usability perception by almost 9 points on average. These changes mitigated the problems of the mobile application identified in the first evaluation.

In general, scores achieved in the final Likert-style questionnaire were also high (see Table 9). Several factors were measured, from general satisfaction, to the usefulness of the assistance instructions obtaining the agreement of the participants. Only statement 4 ("sensors do not cause inconvenience when doing the experiment") had an average value lower than 4. This was mainly due to the wires of the sensors employed in the experimentation and the need to write during Task 2. Maybe using wireless sensors would enhance this value.

We also have to mention that User 14b did not answer Question 3 ("I am able to make experiments by using the DAFIESKU system without being aided by technical staff"). As she mentioned during her interview, she did not think she was going to perform experiments with users during her career. Therefore, she preferred to leave that question unanswered as she could not imagine the need to use a system like DAFIESKU during her activities.

In general, during the interviews, participants declared the DAFIESKU system to be easy to use, both as a researcher and as an experimental subject. Bearing in mind they were using it without prior presentation of the system, it would appear they were able to carry out the requested activities in an intuitive manner. They also thought that if they were requested to use DAFIESKU again, they would easily remember its functionalities.

Although their first language was either Spanish or Basque, almost all of them tested the English version of the application for researchers. There was only one exception:

Table 8: Scores achieved in SUS questionnaires in the second experiment.

User	Researcher	Participant
User 01b	72.5	95
User 02b	92.5	95
User 03b	72.5	67.5
User 04b	77.5	52.5
User 05b	85	80
User 06b	85	80
User 07b	70	75
User 08b	87.5	75
User 09b	85	67.5
User 10b	82.5	90
User 11b	70	65
User 12b	75	85
User 13b	82.5	90
User 14b	92.5	95
User 15b	87.5	85
User 16b	100	100
User 17b	62.5	62.5
Mean	81.18	80
STD	9.81	13.64

User 12b tested the Spanish version of the application during Task 1. No problems were found and the aforementioned user tested the English version of DAFIESKU during Task 3.

During the interviews several comments arose that will guide us when making new versions of the DAFIESKU system. In what follows we will set out the most relevant of these.

For example, possible errors should be clearly explained to participants in the current tablet app. During the experimentation, in several cases (Users 06b, 09b, and 10b) the Bluetooth connection between the tablet and the BITalino was broken and the message "Unfortunately, Dafiesku has stopped" appeared on the tablet screen (see Figure 9). This message did not give any information about what had happened. The way to solve this problem was suggested by the person in charge of the experimentation and consisted in restarting the tablet. Afterwards, the participants had to start Task 2 again. The error only happened once to each of them and at the second attempt they were able to complete Task 2.

When using wired sensors, several participants (Users 03b, 11b, and 16b) suggested using colors to distinguish them, as they initially thought there were no differences between them. But in general, and when completing a Likert-style questionnaire, the participants thought that wireless sensors would make experimenting with physiological signals more comfortable. They also suggested including more commercial equipment in future DAFIESKU versions to get physiological information, as only BITalino and eHealth equipment were considered in the prototype.

Several participants found this system adequate for monitoring health issues (Users 03b, 10b, 11b, 12b, and 17b). In some cases, they also mentioned other applications such as sports (Users 10b and 11b).

Other enhancements suggested by several participants (Users 10b and 14b) were related to existing on/off buttons in the experiment: if at maximum one had to be active (on), activating one should deactivate the previous one. Thereby, the effort needed (measured in operations to be carried out by the users) would be lower (only one click needed instead of one click to deactivate and another one to activate the new one).

Even the *Next* and *Back* buttons were found useful to check whether participants were wearing the sensors correctly (e.g., User 14b). There was no total agreement about maintaining the current version or changing it. For example, User 05b suggested making a vertical scroll instead of *Next* and *Back* buttons. Another option was mentioned by User 13b, who suggested showing a *next* button instead of a *back* button in Figure 7(e) when a sensor was checked and no errors were found.

Other suggestions were related to including new functionalities within the DAFIESKU system. For example, Users 08b and 16b suggested including statistical information within the researcher part, maybe by also using graphics, to analyze the evolution of the participants in the experiments. User 08b also suggested including the option of taking and sending photos in order to let the participants show how they completed the tasks and allow the researchers to check the results achieved by the participants. User 16b was of the opinion that in future versions the researchers would have to design experiments by using a tablet, mobile, or similar devices, instead of personal computers. She suggested working on this idea for DAFIESKU.

In Table 10, lessons learned in order to make new versions of the systems are highlighted, relating to both the server-side application and the mobile application.

8. Conclusions and Future Work

In this paper, we propose a system aimed at remotely recording physiological data of participants without the help of research assistants. It has been developed by the usual iterative development cycle proposed on User-Centered Design [13]. Making prototypes helps solutions to be put into practice in a fast and economical way. Moreover, evaluating developed prototypes serves to acquire feedback from potential system users. This helps developers to think about the issues raised by users while using the system. Two iterations of the system (versions 1.0 and 2.0) are presented and evaluated in this paper. The system was developed to configure data acquisition, perform data collection, and finally store data in a friendly format for the researcher.

Two case studies were carried out to test the adequacy of the proposed approach and the usability of the software. The first evaluation with 5 participants obtained promising results and raised interesting issues, concerning, among others, the lack of support for guiding the participants. Afterwards, a new iteration of the software was developed to address these issues and include enhancements suggested by the participants. Thus, a second case study with 17 participants,

Table 9: Scores achieved in the final Likert-style questionnaire.

	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8
User 01b	5	5	5	4	3	5	5	5
User 02b	4	5	5	1	5	4	5	5
User 03b	3	4	4	3	5	4	4	3
User 04b	3	3	4	4	5	5	5	3
User 05b	5	4	4	2	4	5	4	5
User 06b	4	5	5	3	4	4	4	4
User 07b	4	4	4	2	4	3	4	4
User 08b	4	4	5	4	5	5	4	4
User 09b	4	4	5	2	4	5	4	3
User 10b	5	5	5	4	5	4	3	5
User 11b	4	4	4	4	4	5	5	4
User 12b	4	5	4	2	4	5	5	5
User 13b	5	4	5	2	4	5	5	5
User 14b	5	5		5	5	4	5	5
User 15b	5	5	4	3	5	5	5	4
User 16b	5	5	5	5	5	5	5	5
User 17b	3	3	3	3	4	5	5	3
Mean	4.24	4.35	4.44	3.12	4.41	4.59	4.53	4.24
STD	0.7	0.63	0.63	1.2	0.63	0.63	0.63	0.79
N	17	17	16	17	17	17	17	17

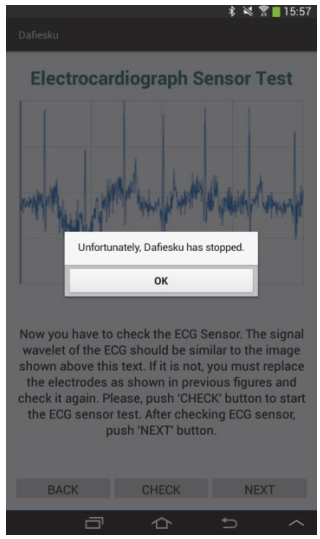


Figure 9: Error screen of the mobile application.

including a student in psychology and other graduates apart from computing, was conducted to measure whether the usability of the system had increased or not.

The results of the second evaluation show an improvement in the mobile application usability in almost 9 points and a slight enhancement of 3 points in the web application on the SUS scale [39]. Therefore the changes made in the mobile side of DAFIESKU were useful to increase the usability perceived by the participants. New feedback and suggestions were collected. Furthermore, the interviews and questionnaires carried out with the participants pointed out that the guidance level required for physiological data acquisition was achieved by the system.

Thanks to the second evaluation, we have identified new features and the need for assistance screens to be added in future versions of DAFIESKU. The next evaluation step should be performed in the homes of a group of volunteers, thereby allowing a full deployment of the system in an uncontrolled environment through the use of participants' personal mobile devices. We also plan to provide the researchers with

Table 10: Lessons learned after case study 2.

#	Server-side application
(1)	Changes made from previous version enhance usability scores.
(2)	Changes made from previous version decrease the time needed to define an experiment.
(3)	It is quite easy to use.
(4)	Include statistical information to analyze participants during the experiments.
(5)	Evaluate the need of collecting multimedia material (for example, photos, videos, and sounds).
(6)	Migrate the system to be portable on a tablet, smart-phone, or similar devices.
#	Mobile application
(1)	Changes made from previous version enhance usability scores.
(2)	Changes made from previous version decrease the time needed to complete the task related to the mobile app.
(3)	It is quite comfortable and easy to use.
(4)	Provide a better explanation of possible errors that may appear.
(5)	Distinguish more clearly which sensors are to be used in experiments (for example, by using colors).
(6)	Try other wireless sensor sets.
(7)	Give the option to personalize the interface to each user's preferences.

more resources to customize their experiments, such as the possibility of editing the text and images that appear in the mobile application assistance screens.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work has been partially funded by the Ministry of Economy and Competitiveness of the Spanish Government and by the European Regional Development Fund (Project TIN2014-52665-C2-1-R) and by the Department of Education, Universities, and Research of the Basque Government under Grant IT980-16. The last three authors belong to the Basque Advanced Informatics Laboratory (BAILab), Grant UFI11/45, partially supported by the University of the Basque Country (UPV/EHU). The authors would like to acknowledge all the participants who took part in the studies.

References

[1] The Encyclopedia of Human-Computer Interaction, 2nd Ed., https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed.

[2] J. A. Jacko, *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, CRC Press, 3rd edition, 2012.

[3] K. Blashkin and P. Isaias, *Emerging Research and Trends in Interactivity and the Human-Computer Interface*, IGI Global, 2014.

[4] J. Holland, *Wearable Technology and Mobile Innovations for Next-Generation Education*, IGI Global, 2016.

[5] UCI Machine Learning Repository, http://archive.ics.uci.edu/ml/.

[6] A. L. Goldberger, L. A. Amaral, L. Glass et al., "PhysioBank, physioToolkit, and physioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–E220, 2000.

[7] H. Blockeel and J. Vanschoren, Experiment databases: towards an improved experimental methodology in machine learning, *Knowledge Discovery in Databases: PKDD 2007*, Springer, Berlin Heidelberg, pp. 6-17, 2007.

[8] D. J. Albers, N. Elhadad, E. Tabak, A. Perotte, and G. Hripcsak, "Dynamical phenotyping: Using temporal analysis of clinically collected physiologic data to stratify populations," *PLoS ONE*, vol. 9, no. 6, Article ID e96443, 2014.

[9] X. Valencia, *A Web Transcoding Framework Based on User Behaviour Evaluation [Ph.D. thesis]*, University of the Basque Country (UPV/EHU, Faculty of Informatics), 2017.

[10] M. Vigo and S. Harper, "Real-time detection of navigation problems on the World 'Wild' Web," *International Journal of Human-Computer Studies*, vol. 101, pp. 1–9, 2017.

[11] UCAmI 2016 Conference Web site: http://mami.uclm.es/ucami-2016/.

[12] E. Sarasua, M. Simón, B. Gamecho, E. Larraza-Mendiluze, and N. Garay-Vitoria, "Physiological data acquisition system based on mobile computing," in *Ubiquitous Computing and Ambient Intelligence*, C. García, P. Caballero-Gil, M. Burmester, and A. Quesada-Arencibia, Eds., vol. 10070 of *LNCS*, pp. 46–51, 2016.

[13] "IDEO: Design Kit: The Field Guide to Human-Centered Design," 2015, https://www.ideo.com/work/human-centered-design-toolkit/.

[14] F. Paternò, "Tools for remote web usability evaluation," HCI International, pp. 828-832, 2003.

[15] F. Christos, K. Christos, P. Eleftherios, T. Nikolaos, and A. Nikolaos, "Remote usability evaluation methods and tools: A survey," in *Proceedings of the 11th Panhellenic Conference in Informatics (PCI*, pp. 151–162, 2007.

[16] X. Valencia, J. E. Pérez, U. Muñoz, M. Arrue, and J. Abascal, "Assisted interaction data analysis of web-based user studies," in *Human-Computer Interaction - INTERACT*, J. Abascal, S. Diniz Junqueira Barbosa, M. Fetter, T. Gross, P. Palanque, and M. Winkler, Eds., vol. 9296 of *LNCS*, pp. 1–19, 2015.

[17] N. Aharony, W. Pan, C. Ip, I. Khayal, and A. Pentland, "Social fMRI: investigating and shaping social mechanisms in the real world," *Pervasive and Mobile Computing*, vol. 7, no. 6, pp. 643–659, 2011.

[18] Funf in a box: http://inabox.funf.org/.
[19] D. Ferreira, *Aware: A Mobile Context Instrumentation Middleware to Collaboratively Understand Human Behavior [Ph.D. thesis]*, University of Oulu, Faculty of Technology, 2013.
[20] H. P. Da Silva, A. Lourenço, A. Fred, and R. Martins, "BIT: biosignal igniter toolkit," *Computer Methods and Programs in Biomedicine*, vol. 115, no. 1, pp. 20–32, 2014.
[21] BITalino: http://www.bitalino.com/.
[22] O. Banos, C. Villalonga, M. Damas, P. Gloesekoetter, H. Pomares, and I. Rojas, "PhysioDroid: combining wearable health sensors and mobile devices for a ubiquitous, continuous, and personal monitoring," *The Scientific World Journal*, vol. 2014, Article ID 490824, 2014.
[23] J. A. Burke, D. Estrin, M. Hansen et al., "Participatory sensing. world sensor web workshop," *ACM Sensys*, 2006.
[24] B. Guo, Z. Yu, X. Zhou, and D. Zhang, "From participatory sensing to Mobile Crowd Sensing," in *Proceedings of the 2014 IEEE International Conference on Pervasive Computing and Communication Workshops, PERCOM WORKSHOPS 2014*, pp. 593–598, March 2014.
[25] B. Resch, M. Sudmanns, G. Sagl, A. Summa, P. Zeile, and J. Exner, "Crowdsourcing physiological conditions and subjective emotions by coupling technical and human mobile sensors," *GI_Forum*, vol. 1, pp. 514–524, 2015.
[26] X. Hu, T. H. S. Chu, H. C. B. Chan, and V. C. M. Leung, "Vita: a crowdsensing-oriented mobile cyber-physical system," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 1, pp. 148–165, 2013.
[27] R. W. Picard, "Future affective technology for autism and emotion communication," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3575–3584, 2009.
[28] Biosignalsplux: http://biosignalsplux.com/index.php/en/.
[29] Shimmer: http://www.shimmersensing.com/.
[30] A. Garzo and N. Garay-Vitoria, "Ethical issues for user involvement in technological research projects: Directives and recommendations," *Contemporary Ethical Issues in Engineering*, pp. 251–269, 2015.
[31] M. Folk, G. Heber, Q. Koziol, E. Pourmal, and D. Robinson, "An overview of the HDF5 technology suite and its applications," in *Proceedings of the EDBT/ICDT 2011 Workshop on Array Databases, (AD'11)*, pp. 36–47, March 2011.
[32] JSON: JavaScript Object Notation, http://json.org/.
[33] Android Studio: Official Android IDE, https://developer.android.com/studio/index.html.
[34] "Butter Knife: Field and method binding for Android views," http://jakewharton.github.io/butterknife/.
[35] "Graph View: Open source graph plotting library for Android," http://www.android-graphview.org.
[36] "Android: Bluetooth API," https://developer.android.com/guide/topics/connectivity/bluetooth.html.
[37] "Android: HttpURLConnection," https://developer.android.com/reference/java/net/HttpURLConnection.html.
[38] Retrofit: A type-safe HTTP client for Android and Java, http://square.github.io/retrofit/.
[39] J. Brooke, "SUS: a retrospective," *Journal of Usability Studies*, vol. 8, no. 2, pp. 29–40, 2013.
[40] J. Nielsen, "How Many Test Users in a Usability Study?" 2012, https://www.nngroup.com/articles/how-many-test-users/.
[41] H. P. Da Silva, A. Fred, and R. Martins, "Biosignals for everyone," *IEEE Pervasive Computing*, vol. 13, no. 4, pp. 64–71, 2014.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 4714328, 15 pages
https://doi.org/10.1155/2017/4714328

Research Article

Exploiting Awareness for the Development of Collaborative Rehabilitation Systems

Miguel A. Teruel, Elena Navarro, and Pascual González

LoUISE Research Group, Computing Systems Department, University of Castilla-La Mancha, Albacete, Spain

Correspondence should be addressed to Elena Navarro; elena.navarro@uclm.es

Received 27 June 2017; Accepted 7 September 2017; Published 17 October 2017

Academic Editor: Pino Caballero-Gil

Physical and cognitive rehabilitation is usually a challenging activity as people with any kind of deficit have to carry out tasks that are difficult due to their damaged abilities. Moreover, such difficulties become even harder while they have to work at home in an isolated manner. Therefore, the development of collaborative rehabilitation systems emerges as one of the best alternatives to mitigate such isolation and turn a difficult task into a challenging and stimulating one. As any other collaborative system, the need of being aware of other participants (their actions, locations, status, etc.) is paramount to achieve a proper collaborative experience. This awareness should be provided by using those feedback stimuli more appropriately according to the physical and cognitive abilities of the patients. This has led us to define an awareness interpretation for collaborative cognitive and physical systems. This has been defined by extending an existing proposal that has been already applied to the collaborative games field. Furthermore, in order to put this interpretation into practice, a case study based on an association image-writing rehabilitation pattern is presented illustrating how this cognitive rehabilitation task has been extended with collaborative features and enriched with awareness information.

1. Introduction

Health care systems are one of the main concerns of both developed and developing countries. To provide their citizens with a proper treatment becomes a must as health is usually considered one of the main pillars of their quality of life. However, offering such services at reasonable costs is not a trivial issue. In this context eHealth [1] has emerged as a proper alternative that facilitates the exploitation of technology to provide both patients and practitioners with solutions according to their specifics needs. According to Black et al. [2] eHealth technologies can be categorized into three main areas: "(1) storing, managing, and transmission of data; (2) clinical decision support; and (3) facilitating care from a distance." Among these areas, the third one, namely, *telerehabilitation* [3], can be considered of special interest because of the important benefits it can offer from the point of view of both patients and policymakers. First, it offers healthcare to patients at their home by using computing technologies, telecommunications, and so forth so that patients with mobility problems do not depend on their relatives to carry out their rehabilitation. Second, policymakers are able to provide more patients with rehabilitation services at reasonable costs [4].

While analyzing the literature of available rehabilitation systems, several technological challenges were detected that must be addressed. One of them is related to the diversity of deficits and diseases a patient may suffer. This means that a telerehabilitation system should provide facilities to design *bespoke therapies,* that is, facilities to adapt the therapy to patients' abilities, disabilities, and needs.

In addition, as aforementioned, these systems are thought to deploy the rehabilitation at patients' home, limiting their interaction with their therapists and with other patients. This means that these systems may hinder the rehabilitation process because of the isolation feeling they may cause. To mitigate such feeling, they may be turned into *collaborative rehabilitation systems* exploiting the concept of *virtual rehabilitation rooms* where they can collaborate with other peers, an important feature that recent studies, such as [5],

have identified as key. For the design of such rooms, an important concept that must be considered from the very beginning is *awareness* [6]; that is, patients must be aware of who is in the virtual room, what their peers are doing, where they are, and so forth. Awareness is a concept that has been traditionally used and exploited in the development of Computer Supported Cooperative Work (CSCW) systems [7]. However, *virtual rehabilitation rooms* are not usual CSCW systems, but their users have their abilities hampered by some kind of trauma, congenital problem, and so forth. Moreover, their development is also highly complex, as virtual reality interfaces, haptic devices, auditory devices, and so forth are used to facilitate users' interaction and to control properly the therapy process. That is, they may be used as different communication channels to provide those feedback stimuli more appropriately according to the physical and cognitive abilities of the patients.

Different awareness interpretations have been defined so far that focus on the identification of awareness for different types of systems. However, none of them focuses on telerehabilitation systems and, thus, they do not consider important characteristics as the users' deficits. In order to address this problem, in this work a new awareness interpretation is presented. It has been defined thanks to our experience of developing rehabilitation systems during the latest years in different domains such as acquired brain injury or children with special education needs, as well as our proposals with CSCW systems. It has not been defined from scratch but extending our previous proposal in a very challenging and demanding domain: video-games development. A set of guidelines is also presented to describe how to put it into practice, as well as a case study to exemplify each awareness element identified in this new interpretation.

This paper has been structured as follows. After this introduction, Section 2 describes the related work. Then, Section 3 describes our proposal, an awareness interpretation for physical and cognitive rehabilitation systems, as well as a set of guidelines for its application. Section 4 illustrates a case study of how this proposal may be applied to design a virtual rehabilitation room. Finally, the conclusions and future work are presented in Section 5.

2. Related Works

The use of Information and Communication Technology advanced (ICT) proposal in the rehabilitation field is not a new approach. As Brennan et al. stated [8], the first solutions of *telerehabilitation* systems date back to almost forty years ago. But, although nowadays it is possible to find different commercial proposals, such as [9, 10], the problem has not been yet solved. The great majority of those commercial applications exhibit several limitations that motivate new researches in this area.

During these forty years of use of ICT in the rehabilitation field, several challenges have been addressed to make the most of the technology. For instance, initial proposals in the area focused on facilitating the communication between therapists and patients, such as [11] who proposed a closed-circuit television to simulate the remote communication. Nowadays, many proposals have revolved around providing new types and complex rehabilitation tasks. For instance, it is possible to find proposals that offer therapies for the treatment of physical [12, 13] and cognitive diseases [14, 15]. Moreover, because of the great diversity of diseases and deficits a patient may suffer, it is difficult to design a general solution applicable to every patient. Other factors [16] such as age, education, and experience with technology must be also considered in the design of these telerehabilitation systems. The availability of systems that therapists may use for the design of bespoke therapies emerges as a solution to address the aforementioned concerns.

Although most of these solutions have been designed to implement specific therapies [17], some of them also enable therapists to adapt some features to the patient's skills [18, 19] or create their own therapies from scratch. Pirovano et al. [20] make use of rehabilitation games and a fuzzy system to adapt such games according to the players' performance. They also offer certain configurability to control the physical stress of the patient. All these proposals that offer therapies adapted to the patient can improve some relevant aspects, such as user motivation and engagement [21].

Along with the need of designing personalized therapies for each specific disease, the use of multisensory feedback is also relevant for improving the rehabilitation of a specific patient as it facilitates a high number of communication channels between patient and system. Even though there are few proposals about the use of multisensory feedback in the area of rehabilitation [19], its inclusion improves the ecological validity of the therapies as the most appropriate communication channel may be selected depending on the specific patient's features. Usually, as Gutiérrez et al. [22] noticed, virtual reality (VR) environments offer a proper solution because of their capabilities to use different communication channels such as visual, aural, and haptic one. Moreover, the exploitation of different channels also helps to enhance how realistic a virtual environment is perceived by the user [10]. For example, the use of haptic sense could enable the transmission of some alerts or information for improving the completion of the task requested [23].

Furthermore, as Cranen et al. claim [24] some of the main advantages of using ICT in the rehabilitation process is to reduce the number of visits to the care center, increase the time available for the therapies, and decrease the treatment costs. However, they also identify some problems resulting from this teletreatment. Namely, patients miss the presence of the therapist and are less motivated when they need to carry out complex exercises. Patients also remark another problem related to their social isolation. Therefore, an advantage as a reduced number of visits to the care center may become a problem, because the direct interaction with therapists and with other patients is also reduced.

In order to mitigate these isolation problems, some authors offer video conference solution [8, 25] and others [26] integrate some social networks in their telehealth systems for improving the interpersonal communication. However, few proposals offer social interaction through some collaborative features that enable patients to collaborate while they carry out a specific therapy [21, 27]. Moreover, similarly to real

environment, other people could play a relevant social role in such collaborative therapies: *observers* and *motivators*. Observers are people who are not doing the therapy but provide some social interaction with the patient. Not only do motivators observe the therapy but they can also cheer the patients up by using their voice or making some gestures that patients could perceive. Usually, these observers and motivators are the relatives of the patient, or some specialists.

The introduction of collaborative features entails the use of concepts already applied in the design of Computer Supported Collaborative Work (CSCW) or groupware systems [28]. One of the main CSCW concepts is *awareness* that has been defined as "the up-to-the-moment understanding of another person's interaction within a shared workspace" [6]. In a previous work [29] the available awareness interpretations (collaboration awareness, situation awareness, workspace awareness, context awareness, social awareness, and so on) were analyzed concluding that it is not possible to cover all the features of modern complex collaborative systems by using just only one of them. To solve such problem a thematic analysis was conducted that led to the definition of *Gamespace Awareness* that integrates the existing proposals in order to guide in the specification of the awareness of one of the most complex collaborative systems: collaborative video-games. In addition, in the telerehabilitation domain some other elements should be taken into account, such as how to provide users with the necessary feedback while they are carrying out their rehabilitation tasks. Therefore, in this work, based on these needs already identified in [30] and our previous proposal [29], we propose an awareness interpretation to be used in the design of collaborative physical and cognitive telerehabilitation systems, identifying both which awareness elements may be of interest and which stimuli may be used to interact with patients making the most of their rehabilitation process.

3. Awareness Interpretation for Physical and Cognitive Rehabilitation Systems

The awareness interpretation developed to deal with the awareness features of telerehabilitation systems for physical and cognitive therapies is presented next. This interpretation has been defined by adapting Gamespace Awareness [29] to make it suitable for this kind of systems.

Table 1 presents the elements identified for this awareness interpretation. They are categorized into four different concerns related to either the temporal stages the awareness is related to (present, past, and future) or those related to social and group dynamics. Within each concern, each awareness element is classified depending on what awareness information they are providing. For instance, within the *present* concern, the awareness elements deal with *who* is participating, *what* and *where* they do anything, and *how* to do it. Aimed at helping designers to identify the awareness needs of a rehabilitation system, a set of questions have been defined along with these awareness elements. For instance, consider a remote physical rehabilitation system whose patients have to work in a collaborative manner. In order to perform such collaboration, they must be aware of *who* is available to collaborate with while performing the rehabilitation session, *what* the other patients are doing to coordinate their actions properly, *where* they are located in their own space, and *how* to interact with them. This awareness information, which is easily perceived in a nonremote and computerless rehabilitation environments, is key when collaborating remotely. Consequently, this new awareness interpretation aims at making remote rehabilitation as straightforward and fruitful as face to face rehabilitation.

As shown in Table 1, for each awareness element, it has been also defined which feedback stimuli could be used to provide participants with the required awareness, that is, *visual* (by means of a computer screen or a virtual reality headset), *aural* (by playing sound or audio messages), or *haptic* (by receiving vibrations on different parts of the body). Furthermore, when dealing with disabled people, providing awareness information through different stimuli is paramount since certain stimuli may not be received or understood by disabled users. For example, when a deaf patient is using a telerehabilitation system, audio messages may be replaced with visual notifications or haptic signals. However, not all awareness elements can be properly provided by using any stimulus. As an example, making the participants aware of the log-in of a new user in the system (element *Identity*) by using haptic feedback would not be practical since it would require codifying each new participant ID as different haptic signals. This would cause that our users were overwhelmed due to the constant vibrations received, and it would be difficult to implement as the number of users increases. Nevertheless, providing the same awareness information through an audio message (e.g., "John is now online") is comprehensible by most of the participants.

Table 1 also presents several examples of how each awareness element could be implemented, that is, how to gather such awareness information and how to provide participants with it. For instance, the *status* element, which is related to the participant's physical and emotional status could be obtained by using either a biometric sensor that gathers physical data such as heat rate or skin conductance [31], or a camera along with an emotional analysis software to analyze participants' emotions [32]. Moreover, this awareness information may be provided by using different stimuli. For instance, if it was required to make a participant aware of other participants' heart rate, this awareness information could be provided by using the three considered stimuli. Firstly, visual stimuli may be easily used by playing animated heartbeats on the screen, thus representing the heart rate of the remote participant. Secondly, aural stimuli can also be considered and implemented by playing heartbeats though audio. Finally, it could be also possible to use haptic stimuli as well, thus emulating heartbeats by sending haptic impulses that the participant will feel on a specific part of his/her body.

3.1. How to Put into Practice the Awareness Interpretation. Once the awareness interpretation has been presented, it will be explained in the following how to put it into practice. For this aim, let us start from the specification of the tasks of a rehabilitation system to be developed. To perform such specification, any requirement engineering specification

Table 1: Awareness elements for physical and cognitive rehabilitation.

Category	Awareness element	Specific questions	Recommended feedback stimuli*			Implementation examples	
			V	A	H	Gathering	Providing
			Present				
Who	Presence	Is anyone in the system?	X	X	X	Motion sensing	Notification of participant login
	Identity	Who is participating? Who does this avatar belong to? Who is available to collaborate with?	X	X		Face recognition	Recognized participant ID
	Authorship	Who is doing that?	X	X		Motion sensing	Notification of current action's authorship
What	Task	What are they doing? What is the difficulty of this task?	X	X	X	Motion sensing	Visualization, hearing, and haptic feedback of remote action
	Goal	What goal is that task part of?	X	X		Manual input	Notification of current goal
	Object	What object are they working on? What object can I work with?	X	X		Motion sensing with object recognition	Visual notification of currently used objects
	Status	What are the participants' status? What are their feelings? What is the objects' status?	X	X	X	Biometric sensor/emotion detection	Visualization, hearing, and haptic feedback of participant's status and feelings, as well as object's status
	Disabilities	What are the participant disabilities? What are they not able to do because of such disabilities?	X	X		Manual input	Visual warnings about their disabilities
	Perception	What are the other participants perceiving? (Looking, touching, hearing, etc.)	X	X	X	Head-mounted camera with mic and motion sensing	Visualization, hearing, and haptic feedback of other participants' perception
Where	Location	Where are the participants/avatars participating? Is it a physical or virtual location?	X			Motion sensing/GPS	Map locations
	Gaze	Where are the participants looking/facing?	X			Eye tracking	Visualization of other participants gaze
	Position	Where is an object? How near is it?	X	X	X	Object and proximity detection	Visualization, hearing, and haptic feedback of position and nearness of objects
	Mobility	What is the participants' mobility range?	X			Motion sensing	Moving avatars showing participants moving range
How	Device	How do I use a certain device to interact?	X	X		Hardware detection/Manual input	Audio and video instructions

Table 1: Continued.

Category	Awareness element	Specific questions	Recommended feedback stimuli*			Implementation examples	
			V	A	H	Gathering	Providing
			Past				
How	Object history	How did this object come to be in this state?	X			Motion sensing with object recognition	Visual evolution of object state
When	Event history	When did that event happen? How often? Is there any network delay?	X			Motion sensing with object recognition	Timestamp of object changes
Who	Presence history	Who was here and when?	X			Face recognition	Log-ins list
Where	Location history	Where has a participant/avatar been?	X			Motion sensing/GPS	Heat maps of participants' locations
	Position history	Where has an object been?	X			Object and proximity detection	Trajectories of objects' positions
	Mobility history	What were the participants' mobility ranges?	X			Motion detection	Moving avatars showing participants previous moving ranges and date
What	Task history	What has a participant/avatar been doing?	X			Motion sensing	Visual list of past tasks
			Future				
When	Next event	When will the next event happen? How often?	X	X	X	Scheduler	Visualization, hearing, and haptic feedback prior to the event
Who	Next participant	Who will be the next participant?	X	X		Face recognition	Notification of the identity of the new participant to join the session
Where	Next location	Where will a participant/avatar be?	X			GPS-based predictor	Possible next location on map
	Next position	Where will an object be?	X			Scheduler	Alert of new position on map
	Next mobility	What will the participants' mobility ranges be?	X			Motion sensing-based predictor	Moving avatars showing participants future moving range based on current progress
What	Next task	What will happen next?	X	X		Scheduler	Visual and aural alert of next task to do
	Next status	What will the participants/avatar next status be? What will the next status of the object be?	X	X		Biometric sensor-based predictor	Visual and aural warning of participants' dangerous physical state

Table 1: Continued.

Category	Awareness element	Specific questions	Recommended feedback stimuli*			Implementation examples	
			V	A	H	Gathering	Providing
		Social & group dynamics					
Who	Members	Who are the members of my rehabilitation group? Has anybody joined/left the group?	X			Manual input	List of rehabilitation group members
	Other members	Who are the members of the other rehabilitation groups?	X			Manual input	List of all rehabilitation group members on therapist view
	Specialist	Who is the specialist conducting the rehabilitations session? Is he/she following the session?	X	X	X	External log-in	Visualization, hearing, and haptic sensing of the specialist conducting the session
	External participants	Who is participating externally in the rehabilitation session? Are they motivators or just observers?	X	X	X	External log-in	Visualization, hearing, and haptic sensing of remote participants
What	Belonging	What rehabilitation group do I belong to?	X	X		Manual input	Visual or aural notification of current group's name
	Role	What is my role within my group?	X			Manual input	Visualization of participant role
	Others' roles	What are the roles of my group's members?	X			Manual input	Visualization of all participants' role on therapist view
	Group goal	What are the goals of my group?	X	X		Manual input	Notification of current group goal
	External motivation	What are other participants doing to motivate me to carry out my task?	X	X	X	Remote voice and commands	Visualization, hearing, and haptic sensing of remote cheering
How	Inner communication	How should I communicate with each rehabilitation session participant?	X	X		Audio recording	Text and audio instructions of how to communicate with the other participants
	Outer communication	How can I communicate with the therapist/external viewers?	X	X		Audio recording	Text and audio instructions of how to communicate with the therapist/external viewers

*(V) visual, (A) aural, and (H) haptic.

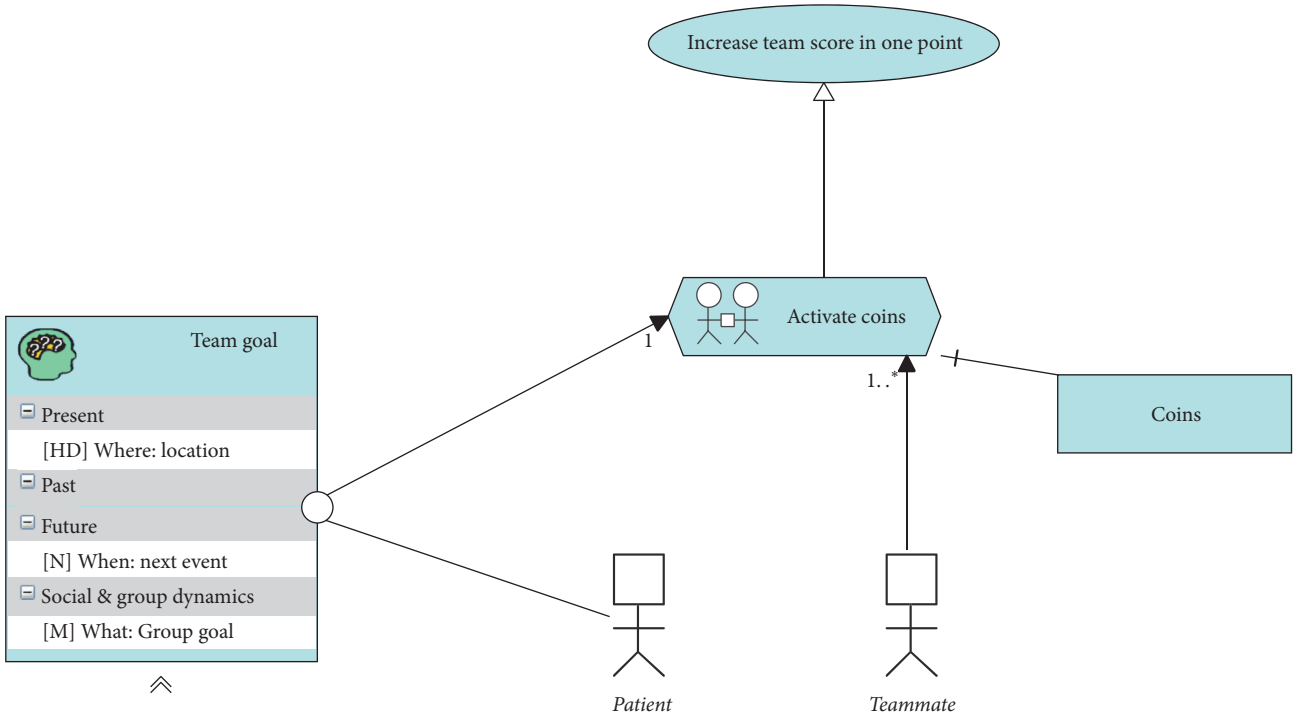


Figure 1: Example of CSRMF diagram specifying a collaborative task with awareness needs. $*$ means N; that is, from 1 to N teammates can be related to activate coins.

technique may be used. However, due to its collaboration and awareness modelling capabilities, we recommend using the Collaborative Systems Requirements Modelling Framework (CSRMF) [33].

Once the tasks of the system have been specified, those awareness features considered necessary to ease the interaction between the user and this system can be defined. For this aim, Table 1 provides a set of questions that will help us to identify the awareness needs of our system. As an example, let us consider the first awareness element on Table 1, *Present-Who-Presence*. Such element includes the *specific question* "Is anyone in the system?" Therefore, to assess whether or not this awareness element should be included among our system's requirements, we would ask ourselves "Do our users need to know if there is anyone in the system?" If so, this awareness element will be included among our system's requirements. In Table 1, *Present-Who-Presence*, such element includes the *specific question* "Is anyone in the system?" Therefore, to assess whether or not this awareness element should be included among our system's requirements, we would ask ourselves "Do our users need to know if there is anyone in the system?" If so, this awareness element will be included among our system's requirements.

Once we are in the design phase of the system, the stimuli to provide each awareness element must be specified. With this aim, Table 1 also includes the *recommended feedback stimuli* column that suggests what stimuli may be used to communicate the awareness information. For instance, the *presence* element could be offered by using visual, aural, or haptic stimuli (or a combination of them). Table 1 has the *recommended feedback stimuli* column that suggests what stimuli may be used to provide the awareness information. For the considered *presence* element, it could be provided by using visual, aural, or haptic stimuli (or a combination of them). Moreover, Table 1 also includes *implementation examples* for each awareness element. In this case, the information to provide the *presence* element may be gathered by employing a motion-sensing device when it detects that there is a new user in the scenario, provided by notifying the log-in of a new participant using a visual message, playing a log-in sound, or sending a haptic signal.

Figure 1 shows a different example of how to model a task and specify its corresponding awareness needs by using CSRMF. This is a collaborative task whose participants, namely, a *player* and his/her *teammates* must *activate* several *coins* in a virtual world to fulfil their goal: *increase the team score one point*. For this aim, it is *mandatory* (M) that such player be aware of what the coin he/she has to activate, that is, *what* the *group goal is*. Moreover, it is *highly desirable* (HD) to know *where* the other players are *located* in the virtual world. Finally, it would be *nice to have* (N) information about *when* the next *event* will happen, that is, when new coins will appear. This task will be implemented in the case study that will be presented in Section 4.4.

4. Case Study

In the previous section, an interpretation aimed at identifying the awareness requirements of a rehabilitation system was presented. Therefore, to put it into practice, this section presents a case study based on a physical-cognitive collaborative rehabilitation exercise. Aimed at making the

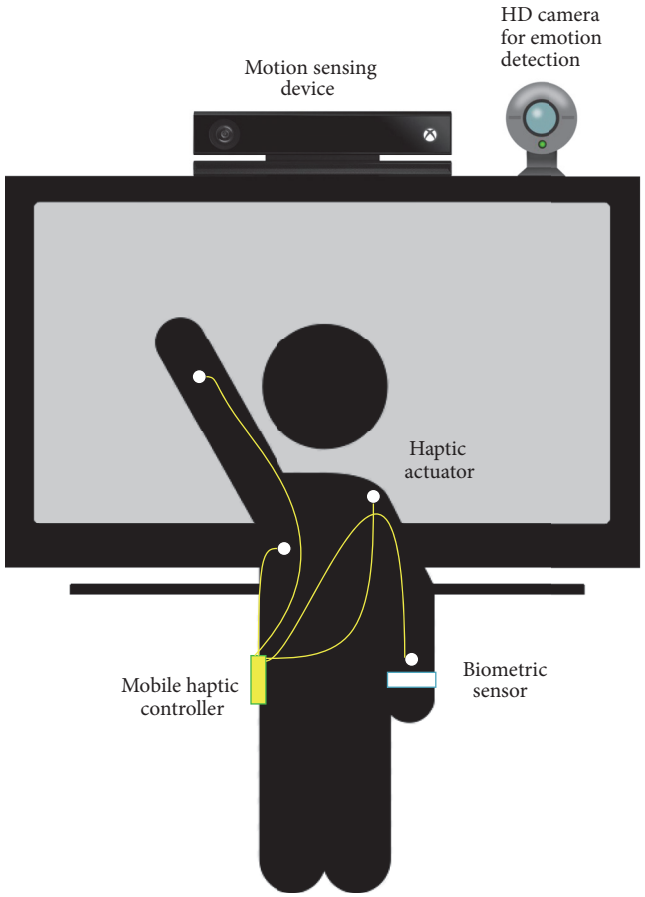


Figure 2: Participants' hardware environment.

performance of such exercise enjoyable, it will be presented as a game whose users will play collaboratively yet remotely grouped in several teams. In this game, patients will be represented as avatars. Hence, to interact with this game when performing the exercise, a motion-sensing device will be used to translate the participants' movements into avatars movements. Figure 2 exemplifies how the hardware environment for this rehabilitation system could be. A motion-sensing device such as Microsoft Kinect [18] will gather the participants' movements. Besides, a high definition camera will capture the users' face and forward it to an emotional analysis SDK [32] that will analyze and interpret their emotional status. A wristband will also interpret their emotional status measuring their heart rate. This physical and emotional information will be sent to the therapist to follow the evolution of the patients. Finally, a wearable haptic controller with several actuators will be used to make the user feel certain events of the game. Hence, 4 actuators will be used, located on both patients' hands (to feel the game objects), chest (to feel heartbeats), and shoulder (to feel encouragement pats).

As far as the cognitive rehabilitation exercise is concerned, it is based on an *association image-writing* rehabilitation pattern. This pattern is defined in [34] as an exercise to improve patients' front executive capability.

Figure 3: Prototype of participants' user interface (patient's avatar located in the center of the interface using a third-person view).

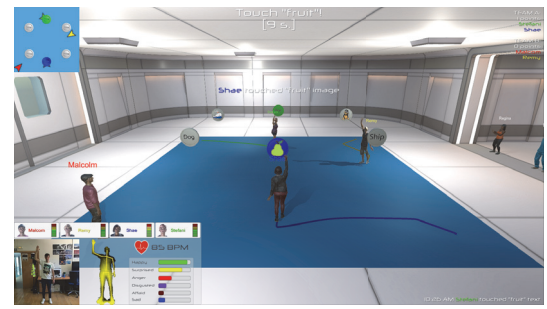


Figure 4: Prototype of therapist's user interface.

Regarding the game itself, the participants' avatars will be in a virtual world where several coins with images and words on them will be scattered around the game field. Therefore, the implementation of the aforementioned rehabilitation pattern consists in asking the teams of participants to find and touch a virtual coin related to a specific concept. As an example, if the system request is to find and touch the "fruit" coin, one participant will have to touch the coin with the image of a pear on it; meanwhile the other does so with the coin showing the word "pear." Figure 3 illustrates two different avatars. The avatar controlled by the current player is trying to activate a coin; meanwhile the green player has already activated the "pear" coin.

Besides, regarding physical rehabilitation, the therapist will be able to place the coins in different locations to encourage the participants' displacements and movements. For instance, if a participant needs an upper limb rehabilitation therapy, the therapist could locate the coins in a high position, so that the participant will have to lift his arm to activate such coin. Moreover, the size of the game field can be customizable according to the movements requirements of the therapy to be performed. For instance, if a patient could not walk, all the coins would be located around the avatar to avoid displacements over the game field.

Furthermore, the therapist will have a different view of the system (see Figure 4) to monitor the therapy execution, enabling her to see each participant view, the participants themselves, as well as their emotional status (obtained by using a facial analysis software).

Finally, external participants (neither patients nor specialists) can participate into the system. Specifically, these

Figure 5: Prototype of observer's user interface.

external participants will adopt the role of *observers* and *motivators*. Hence, observers can watch the rehabilitation session (see Figure 5). Besides, motivators can also cheer the patients up by using their voice that patients may hear during the rehabilitation session. It is worth noting that both observers and motivators, as well as therapists, will be also represented in the game by using avatars to make patients aware of the presence of such participants (see Figure 3).

Aimed at exemplifying how the awareness interpretation presented in Section 3 could be used in the design of such rehabilitation exercise, in the following it is explained how several awareness elements could be implemented and presented to the participants for the recommended feedback stimuli described in Table 1. Notice that several elements could not be implemented in this case study since they are not suitable for this exercise. For instance, in this game every participant has the same responsibilities within his/her team, without having a specific role. Consequently, the *role* and *others' roles* awareness elements were not used.

4.1. Present. The first awareness elements are those related to what is happening in the rehabilitation session at the very moment of its performance, namely, the *present*. Hence, the implemented elements belonging to such concern are the following ones:

(i) *Presence*. Participants have to be aware of the presence of a new participant with whom to collaborate. Following the suggested implementation described in Table 1, the exercise could be implemented according to the following requirements: when a motion-sensing device detects a new participant, the system will provide the already-connected participant (if any) with a *haptic* stimulus indicating the presence of a teammate with whom to perform the rehabilitation session. Therefore, it would not be necessary that a logged-in participant is continuously looking at the screen to know whether there is another user in the virtual rehabilitation room, being warned with a vibration on his wrist when the session is ready. However, this awareness information could also be provided by means of *visual* (by showing the participant's avatar) or *aural* signals to indicate a new log-in into the system.

(ii) *Identity*. Participants have to be aware of which participant is related to which avatar. With this aim, each participant's avatar will have a tag with the participant's name over the head (Figure 6). Moreover, each participant name will have a different color aimed at easing the visual identification of participants.

(iii) *Authorship*: Participants have to be aware of who activated a coin in the virtual world. The requirement would be the following: the color of such coin (*visual* representation) could change to match the color of the participant's avatar. As an example, Figure 3 shows that the pear (image) coin has been touched by the blue player.

(iv) *Task*. The therapist must be aware of what the participants are doing. It could be achieved by means of a *visual* remote view (Figure 4) to observe both what the participants are doing in the virtual room (through their avatars) and the real world (real streamed video of the participants). To reinforce this feedback, an *aural* message such as "Malcom has activated the dog coin" or a *haptic* signal representing a coin activation could be sent to the participants.

(v) *Object*. Participants have to be aware of the coin that they as well as the other participants have activated. For instance, similarly to the *authorship* awareness element, the color of the coins (*visual* representation) could provide awareness information regarding what object they are interacting with. However, different sounds (*aural* representation) could be associated with the coins that will be played when they are touched, thus helping the participants to identify them. For instance, if a participant activated the "dog" coin, a dog bark could be played though the audio system. Finally, in order to make a participant aware of the fact that the target coin has been activated, not only will it change its color according to the participant's color, but also he/she will receive a long haptic signal on the hand he/she activated the coin with.

(vi) *Status*. The therapist has to be aware of both the participants' heart rate and emotional status. To implement this awareness element, the participants' heart rate and emotional status could be presented in a *visual* manner (Figure 4) to the therapist who would be able to interrupt the session or adjust if needed. For instance, if a participant's heart rate was considerably high or he was in a bad emotional state, the therapist could interrupt the session and adapt it according to the participant's needs. The heartbeats of the participants could be coded into *aural* beats or *haptic* vibrations to make them aware of their own heart rate, that is, of their own physical status.

(vii) *Disabilities*. The system has to be aware of any participant's disabilities in order to avoid and/or adapt the stimuli provided. Prior to start a session, the therapist will indicate whether any participant have disabilities to configure the system. For instance, if a participant

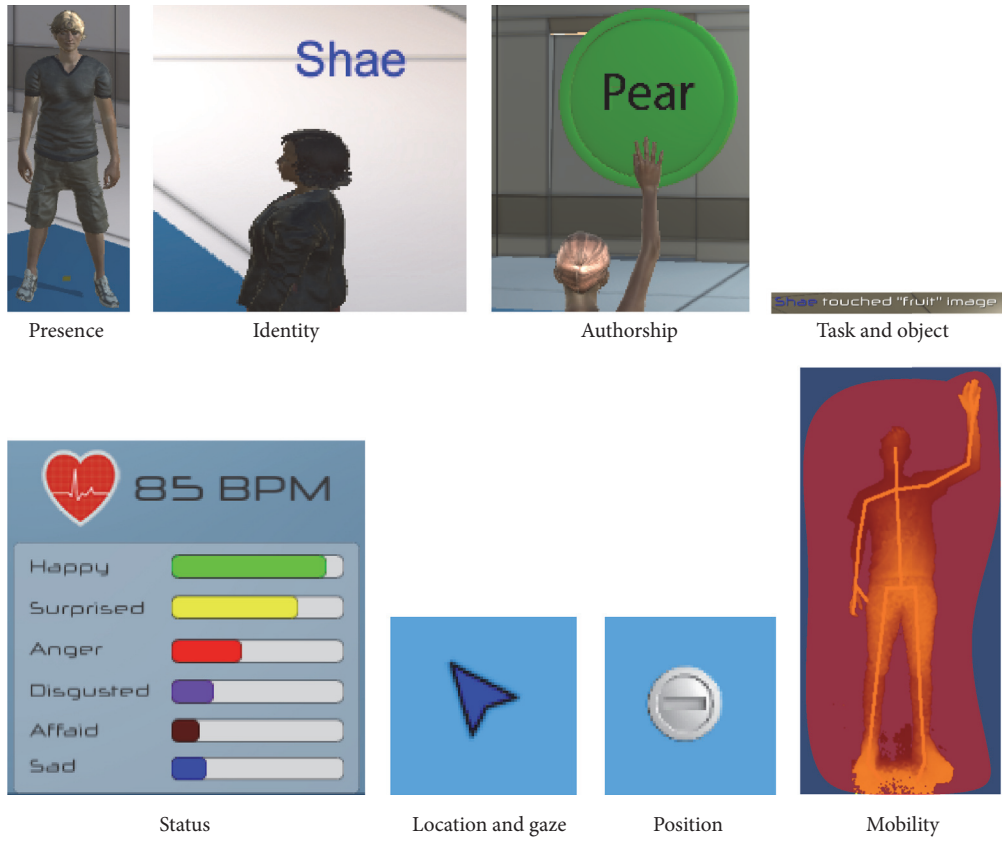


Figure 6: Visual implementation of present awareness element.

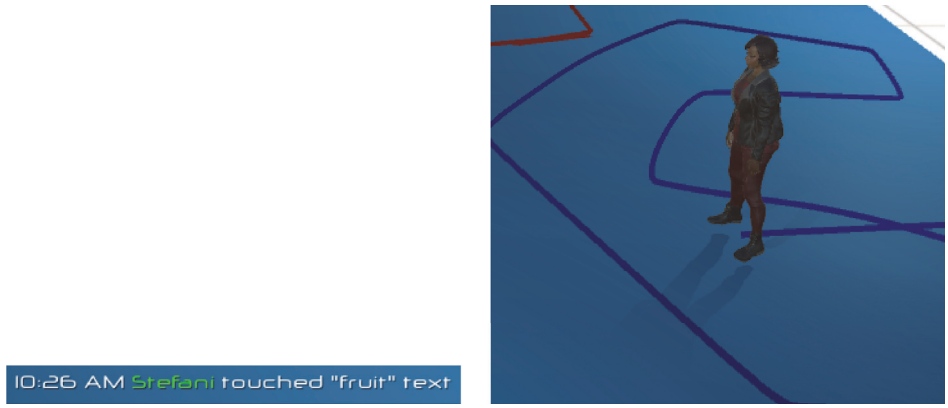
is deaf, audio messages will be presented by using *visual* closed captions. Besides, blind participants may interact with the system by receiving *aural* messages about the coins' location. Moreover, if a participant is unable to walk, all the coins will be located within the participant's reach area.

(viii) *Location*. The participants have to be aware of their locations. The patient will know such location by looking at a *visual* map of the scenario (top-left corner of Figure 3). Moreover, by looking at this map, he may obtain information about the position of the other participants. It is worth noting that this map may represent the participants as arrows with the same color as the participants' name.

(ix) *Gaze*. The direction each participant is facing can be seen in the map by means of the arrows representing the participants (Figure 6). In this sense, the arrow will point towards the direction where the participant is facing. However, if the therapist had to be aware of where the participants are looking to detect cognitive problems, an eye-tracking system could be implemented. For instance, the therapist could see the point of the screen where the participants are looking on the therapist's view in a *visual* manner. Therefore, if a participant was constantly looking at a screen point where no coin or participant was located, a cognitive issue could be identified.

(x) *Position*. Participants must be aware of the position of the coins. Similarly to what happens with the *location* element, participants may know the position of the elements by using this *visual* map (top-left corner of Figure 3). Thanks to this awareness feature, the participants could be aware of the presence of coins positioned behind them that cannot be seen in a 3D third-person view. For a more detailed analysis of the difference between location and position please refer to [29]. To reinforce this feedback, the concept of nearness to the coin could be represented as *aural* messages with a variable pitch or increasing the *haptic* signal depending of such nearness.

(xi) *Mobility*. The therapist has to be aware of what the participants' can reach or where they can move. For example, the therapist may see a *visual* representation of each participant's reach area on a postsession screen (see Figure 4). This image could be generated based on their previous movements. Therefore, the therapist will analyze if the physical rehabilitation process has been successful by measuring a possible enlargement of such reach areas (i.e., the participants are able to reach further with their limbs than before starting the rehabilitation process). This image could be generated based on their previous movements. Therefore, the therapist will analyze if the physical rehabilitation process has been successful by



Event and task history Location history

Figure 7: Visual implementation of past awareness element.

Table 2: Examples of present awareness elements implemented through aural messages.

Awareness element	Aural message example
Presence	[Participant log-in sound]
Identity	Shae has joined the game
Task and object	Malcom has activated the dog coin
Object	[Dog bark]
Status	[Heartbeat]
Position	[Pitch-variant sound]
Device	Walk toward a coin and raise your hand to touch it

measuring a possible enlargement of such reach areas (i.e., the participants are able to reach further with their limbs than before starting the rehabilitation process).

(xii) *Device*. Participants have to be aware of how to interact with the system. For instance, the interaction with the system in this case study will be performed by means of the motion-sensing device. Therefore, if the system detects that a participant was not aware what he has to do, a *visual* or *aural* message would be displayed or played, respectively, thus informing such participant about how to interact (Table 2).

4.2. Past. Once the awareness elements related to the present have been presented, those related to the *past* will complement the participants' awareness with information about facts that happened prior to the current time:

(i) *Object history*. Once the session is finished, a post-game interface could provide detailed information of the different states each coin had had. In other words, this will implement a log detailing the interaction of participants with the game coins.

(ii) *Event history*. The game interface will show the exact time when something happened. This information will be reflected in the actions log located at the bottom-right corner of the interface (see Figure 3).

(iii) *Presence history*: The presence of participants in the system will be logged to be queried by the therapist when required. Hence, the system will record both the log-in and log-out time of each participant. With these data, the therapist could analyze whether or not the participants have been performing their assigned therapies regularly.

(iv) *Location history*. To make both participants and therapist aware of the previous location of a participant's avatar, they will leave a trail behind as they move around the game field (see Figure 7). This trail will have the same color of the participant's name, and the time that it will remain shown can be configured.

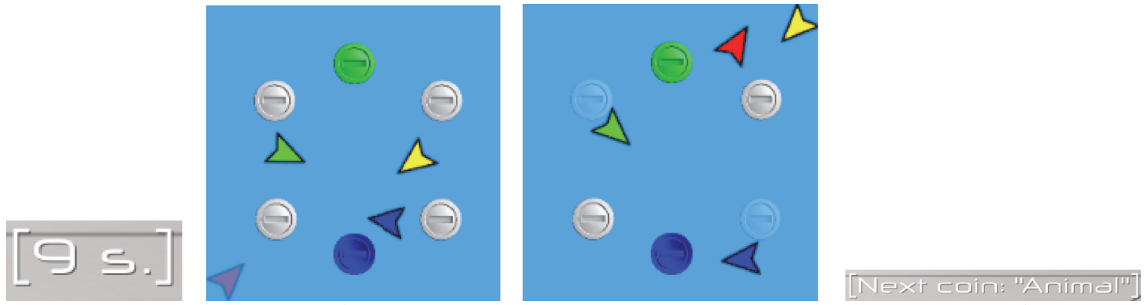
(v) *Position history*. This awareness element will complement *object history* by adding information about where each coin appeared. In this sense, the game log will also include the position of each coin in the map. This could be implemented graphically (indicating the position of the coins by using a map) or textually (just by indication of the coordinates of the position of each coin in meters, departing from the center of the game field).

(vi) *Mobility history*. Closely related to the *mobility* element from the present concern, the mobility history can represent the evolution of the patient's mobility over time. In this sense, it can be analyzed how the patient's mobility range has been evolved from the very beginning of the rehabilitation therapy. Therefore, the therapist can assess whether or not the therapy has been fructiferous for each patient.

(vii) *Task history*. The system will record and show all the tasks performed by the participants. Therefore, a task history will be shown in the bottom-right corner (see Figure 3). It is worth noting that, in order to ease the identification of participants in this list of tasks, each name will be colored according to the participant's color.

Table 3: Examples of future awareness elements implemented through aural messages.

Awareness element	Aural message example
Next event	Next coin in 9 seconds
Next participant	A new participant will join team 1 in 15 seconds
Next mobility	After this session, your right arm mobility could improve 2%
Next task	The next coin will be "Animal." Get ready!
Next status	[Warning sound]



Next event | Next location | Next position | Next task

Figure 8: Visual implementation of future awareness element.

4.3. Future. Next, we will deal with those awareness elements related to the *future*. Such elements make the participants be aware of something that will happen in a near future which could affect the current development of the rehabilitation session somehow. The future awareness elements are the following ones:

(i) *Next event*. In order for the participants to anticipate their actions, this element will provide them with information about what the next event will be. In this example, this element will show the remaining time that a new coin will appear in the game field (see top of Figure 3).

(ii) *Next participant*. During a game, no new participants can join it. However, when a new coin is going to appear, a new participant can join the game, sending an audio message to the current participants (Table 3). Hence, they will be aware that a new participant will join the game soon, as well as what team he/she will be part of.

(iii) *Next location*. Related to the previous awareness element, prior to the appearance of a new participant, his/her starting position can be signaled on the map (Figure 8). In this sense, not only will current participants be aware of an upcoming new participant, but they will also know where he/she will be located initially.

(iv) *Next position*. Similar to the previous element, when a new coin is going to appear, it could be signaled on the map so that participants get ready to touch such coin (Figure 8). Thus, they could anticipate the appearance on this new coin moving towards the new position where it is supposed to appear.

(v) *Next mobility*. Based on *mobility history*, the system could analyze the progression of the patients' mobility over time. Therefore, the system could predict what the future mobility of the user will be if he/she keeps on following the rehabilitation session. This information can be used to motivate patients prior to the performance of a new session by making them aware of what they can achieve during the next sessions.

(vi) *Next task*. The element *next event* was used to make participants aware that a new event will happen. Complementarily, such awareness information can be enriched by adding details of such event. In this example, this information could advance the next task that participants will have to perform, namely, the coin that they will have to touch next.

(vii) *Next status*. This awareness element could be used to make a patient (and the therapist) aware that he will be in a specific emotional or physical state based on past measurement. As an example, if the participant is having a high heart rate in combination with a negative emotional status (sad or disguised), the system could recommend that he/she suspend the rehabilitation session.

4.4. Social and Group Dynamics. Finally, the last awareness elements are related to a nontemporal concern, namely, social and group dynamic. Hence, the elements belonging to this last concern are the following ones:

(i) *Members*. The participants will be able to know who is participating in their team at any time during the rehabilitation session. With this aim, a list of players belonging to his/her own team will be shown on

Table 4: Examples of social and group dynamics awareness elements implemented through aural messages.

Awareness element	Aural message example
Specialist	This session is being supervised by Dr. Smith
Belonging	You will play as a member of team 2
Inner communication	Speak to communicate with your teammates. They will hear you!
Outer communication	You can always communicate with your therapist by using your voice Let him know if you have any problem while performing the therapy

Members and other members | Specialist and external participants | Group goal | Outer communication

Figure 9: Visual implementation of social and group dynamics awareness element.

the participant user interface (see top-right corner of Figure 3). Moreover, to ease the identification of team members, the name of the teammates on this list will be colored with each participant name color.

(ii) *Other members*. As happened with the previous awareness element, a different list will show the members of the rival teams with their corresponding character colors (see top-right corner of Figure 3).

(iii) *Specialist*. The participants will be able to know who is the therapist conducting the session. Moreover, to recreate a real scenario, an avatar representing the therapist will be included in the game interface, also showing the therapist's name (Figure 9).

(iv) *External participants*. During the sessions, two kinds of external participants are allowed to interact, namely, observer and motivators. The former will be able to watch the session; meanwhile the latter can also cheer the participants up by using their voice. Therefore, to make the participants aware of the presence of such external participant, they will be represented as avatars observing or cheering up the session from outside the game field (Figure 9).

(v) *Belonging*. In this example, participants are not allowed to choose the team they belong to since the therapist decides these teams in advance. Because of that reason, the system will make them aware of the team they belong to at the very beginning of the session through an audio message (Table 4).

(vi) *Group goal*. The participants have to be aware of their group goals. For example, considering as goal the coin they will have to touch, it could be communicated by means of both a *visual* message on the screen and an *aural* notification. Thanks to this double awareness, participants would have instant information about the goal as soon as it changes but they would always be able to watch it on the screen (at the top of Figure 3).

(vii) *External motivation*. When remote participants acting as motivators are monitoring the session, they can cheer the participants up. This can be done by using the three considered stimuli. First, they can make their avatars clap so that the participants could see and hear them. Moreover, this motivation may be also transmitted by using the motivators voice so that the patients will hear the encouragement voice messages of the remote participants. Finally, motivators may also send haptic commands. In that sense, patients will receive and feel a haptic pat on their shoulder.

(viii) *Inner communication*. If the therapist decided so, the participants will be able to communicate with their teammate by using their voice. Hence, if this interaction is enabled and the system detected that one participant is not speaking at all, an audio message will inform such participant that his/her voice will be heard by the other team member (Table 4).

(ix) *Outer communication*. Once again, the therapist can decide whether participants can interact verbally with the external participants as well as with the therapist himself/herself. Once again, the system can remind the participants that their voice will be heard by those external participants. Moreover, the therapist will be also able to see a VU meter corresponding to each participant's microphone to be aware visually when a patient is speaking (Figure 9).

Thanks to the implementation of these awareness elements, as well as the others belonging to the present, past, and future concerns, participants will be able to interact with a rehabilitation system like the one presented at the beginning of this section. This will enable collaboration with remote participants, and it will enable the therapist to monitor and adapt the therapies to the participants needs.

5. Conclusions and Future Work

In order to achieve a proper collaboration while interacting with remote participants, they have to be aware of their collaborators (i.e., what they are doing, where they are, etc.). Therefore, dealing with awareness information properly is paramount for the success of a collaborative system. This issue is even more crucial when dealing with rehabilitation systems, whose participants may suffer from any cognitive or physical disability. This turns awareness information into a cornerstone of telerehabilitation systems, where awareness may lead to a suitable interaction. In order to guide the identification of such awareness requirements when developing cognitive and physical rehabilitation systems, an already existing awareness interpretation, namely, Gamespace Awareness [29], has been chosen as the foundation of our proposal. Such awareness interpretation was initially developed to deal with the awareness requirements of collaborative games. In this work, it has been adapted to make it suitable for rehabilitation systems featuring the monitoring of patients by using mobile sensors and providing feedback by using different kinds of stimuli (visual, aural, and haptic). This new interpretation comprises 39 awareness elements (classified into 4 different concerns) that will provide both patients and therapists with the required awareness information to ease collaborative and remote rehabilitation therapies. Furthermore, along with such collection of elements, the awareness interpretation provides a series of specific questions which will help designers and developers of new rehabilitation systems to identify awareness requirements properly. Moreover, we also provide recommendation of which feedback stimuli could be used to provide each awareness element (visual, aural, or haptic), as well as implementation examples for gathering and providing each one. Finally, in order to ease the usage of this awareness element in the development of a new rehabilitation system, a series of guidelines have been provided (Section 3.1).

In order to exemplify the awareness interpretation defined for this type of systems, a case study has been presented (Section 4). It describes several participants while performing a collaborative physical and cognitive rehabilitation exercise. The interaction with the system is performed by means of motion-sensing devices that translate the participants' movements into avatars' movements in the virtual world. Moreover, biometric devices and emotional analysis software are used to make the therapist aware of the patients' physical and emotional status.

It is worth noting that the game presented as case study is just a prototype. Therefore, as a future work, the final game will be implemented. Indeed, to perform this implementation, the postgame user interfaces will be implemented. This functionality will enable therapists to follow the evolution of their patients towards a successful rehabilitation. With the game fully implemented, it will be integrated with our rehabilitation management system to enable us to start the testing of such game by involving real users.

Disclosure

This paper is an extension of Teruel et al. [35].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by Spanish Ministry of Economy, Industry and Competitiveness, State Research Agency/European Regional Development Fund, under Grants Vi-SMARt TIN2016-79100-R and HA-SYMBIOSIS TIN2015-72931-EXP.

References

[1] H. Oh, C. Rizo, and M. Enkin, "What is eHealth (3): a systematic review of published definitions," *Journal of Medical Internet Research*, vol. 7, no. 1, 2005.

[2] A. D. Black, J. Car, C. Pagliari et al., "The impact of ehealth on the quality and safety of health care: a systematic overview," *PLoS Medicine*, vol. 8, no. 1, Article ID e1000387, 2011.

[3] D. M. Brennan, S. Mawson, and S. Brownsell, "Telerehabilitation: enabling the remote delivery of healthcare, rehabilitation, and self management," in *Studies in Health Technology and Informatics*, pp. 231–248, 2015, http://www.ncbi.nlm.nih.gov/pubmed/19592797.

[4] European Commission Information Society and Media, *ICT for Health and i2010: Transforming the European healthcare landscape towards a strategy for ICT for Health*, European Comission, Luxembourg, 2006.

[5] E. Doig, J. Fleming, and P. Kuipers, "Achieving Optimal Functional Outcomes in Community-Based Rehabilitation following Acquired Brain Injury: A Qualitative Investigation of Therapists' Perspectives," *British Journal of Occupational Therapy*, vol. 71, no. 9, pp. 360–370, 2008.

[6] C. Gutwin and S. Greenberg, "A Descriptive Framework of Workspace Awareness for Real-Time Groupware," *Computer Supported Cooperative Work (CSCW)*, vol. 11, no. 3-4, pp. 411–446, 2002.

[7] C. Gutwin and S. Greenberg, "Effects of awareness support on groupware usability," in *Proceedings of the the SIGCHI conference*, pp. 511–518, Los Angeles, California, United States, April 1998.

[8] D. M. Brennan, A. C. Georgeadis, C. R. Baron, and L. M. Barker, "The Effect of Videoconference-Based Telerehabilitation on Story Retelling Performance by Brain-Injured Subjects and Its Implications for Remote Speech-Language Therapy," *Telemedicine Journal and e-Health*, vol. 10, no. 2, pp. 147–154, 2004.

[9] Y. Chang, S. Chen, and J. Huang, "A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor

disabilities," *Research in Developmental Disabilities*, vol. 32, no. 6, pp. 2566–2570, 2011.

[10] H. G. Hoffman, A. Hollander, K. Schroder, S. Rousseau, and T. Furness, "Physically touching and tasting virtual objects enhances the realism of virtual experiences," *Virtual Reality*, vol. 3, no. 4, pp. 226–234, 1998.

[11] R. T. Wertz, N. F. Dronkers, E. Bernstein-Ellis, Y. Shubitowski, R. Elman, and G. K. Shenaut, *Appraisal and Diagnosis of Neurogenic Communication Disorders in Remote Settings, Clinical Aphasiology*, vol. 17, 1987.

[12] G. Pison, *1914–2014: A century of change in the French population pyramid*, vol. 509, Institut National d'Études Démographiques (INED), 2014.

[13] M. Sandlund, E. L. Waterworth, and C. Häger, "Using motion interactive games to promote physical activity and enhance motor performance in children with cerebral palsy," *Developmental Neurorehabilitation*, vol. 14, no. 1, pp. 15–21, 2011.

[14] O. Gervasi, R. Magni, and M. Zampolini, "Nu!RehaVR: virtual reality in neuro tele-rehabilitation of patients with traumatic brain injury and stroke," *Virtual Reality*, vol. 14, no. 2, pp. 131–141, 2010.

[15] A. Cagnin, N. Jelcic, M. Agostini et al., *Feasibility and efficacy of cognitive telerehabilitation in early Alzheimeramp;rsquo;s disease: aamp;nbsp;pilot study, Clinical Interventions in Aging*, 2014.

[16] D. M. Brennan and L. M. Barker, "Human factors in the development and implementation of telerehabilitation systems," *Journal of Telemedicine and Telecare*, vol. 14, no. 2, pp. 55–58, 2008.

[17] I. Dimbwadyo-Terrer, A. de los Reyes-Guzmán, A. Bernal-Sahún et al., "Virtual Reality System Toyra: A New Tool to Assess and Treatment for Upper Limb Motor Impairment in Patients with Spinal Cord Injury," in *Converging Clinical and Engineering Research on Neurorehabilitation*, vol. 1 of *Biosystems & Biorobotics*, pp. 853–858, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[18] G. Kurillo, J. J. Han, A. Nicorici, and R. Bajcsy, "Tele-MFAsT: Kinect-Based Tele-Medicine Tool for Remote Motion and Function Assessment," in *Studies in Health Technology and Informatics*, pp. 215–221, IOS Press, 2014.

[19] T. Lisa, "The use of multisensory environment in acquired brain injury," in *Proceedings of the The 2012 Annual Autumn Conference of the British Association of Neuroscience Nurses*, Edinburgh, Scotland, 2012.

[20] M. Pirovano, R. Mainetti, G. Baud-Bovy, P. L. Lanzi, and N. A. Borghese, "Self-adaptive games for rehabilitation at home," in *Proceedings of the 2012 IEEE Conference on Computational Intelligence and Games (CIG)*, pp. 179–186, Granada, Spain, September 2012.

[21] V. López-Jaquero and F. Montero, "Specifying How to Motivate People in Computer Assisted Rehabilitation," in *Ubiquitous Computing and Ambient Intelligence*, vol. 10069 of *Lecture Notes in Computer Science*, pp. 99–110, Springer International Publishing, Cham, 2016.

[22] M. A. Gutiérrez, F. Vexo, and D. Thalmann, *Stepping into Virtual Reality*, Springer London, London, 2008.

[23] A. D. Cheok, X. Yang, Z. Z. Ying, M. Billinghurst, and H. Kato, "Touch-space: mixed reality game space based on ubiquitous, tangible, and social computing," *Personal and Ubiquitous Computing*, vol. 6, no. 5-6, pp. 430–442, 2002.

[24] K. Cranen, C. H. Drossaert, E. S. Brinkman, A. L. Braakman-Jansen, M. J. IJzerman, and M. M. Vollenbroek-Hutten, "An exploration of chronic pain patients' perceptions of home telerehabilitation services," *Health Expectations*, vol. 15, no. 4, pp. 339–350, 2012.

[25] J. Cason, "A Pilot Telerehabilitation Program: Delivering Early Intervention Services to Rural Families," *International Journal of Telerehabilitation*, vol. 1, no. 1, 2009.

[26] Y. Huang and Y. Hsu, "Social networking-based personal home telehealth system: A pilot study," *Journal of Clinical Gerontology and Geriatrics*, vol. 5, no. 4, pp. 132–139, 2014.

[27] R. Llorens, M. Alcaniz, M. D. Navarro, J. Ferri, and E. Noe, "Self-awareness rehabilitation through a multi-touch virtual game board after acquired brain injury," in *Proceedings of the 2013 International Conference on Virtual Rehabilitation (ICVR)*, pp. 134–138, Philadelphia, PA, USA, August 2013.

[28] J. Grudin and S. Poltrock, "Taxonomy and theory in computer supported cooperative work," in *The Oxford Handbook of Organizational Psychology*, S. W. J. Kozlowski, Ed., pp. 1323–1348, Oxford University Press, 2012.

[29] M. A. Teruel, E. Navarro, P. González, V. López-Jaquero, and F. Montero, "Applying thematic analysis to define an awareness interpretation for collaborative computer games," *Information and Software Technology*, vol. 74, pp. 17–44, 2016.

[30] M. A. Teruel, E. Navarro, and P. González, "Towards an Awareness Interpretation for Physical and Cognitive Rehabilitation Systems," in *Ubiquitous Computing and Ambient Intelligence*, vol. 10069 of *Lecture Notes in Computer Science*, pp. 121–132, Springer International Publishing, Cham, 2016.

[31] J. Cusveller, C. Gerritsen, and J. de Man, "Evoking and Measuring Arousal in Game Settings," in *Games for Training, Education, Health and Sports*, vol. 8395 of *Lecture Notes in Computer Science*, pp. 165–174, Springer International Publishing, Cham, 2014.

[32] H. Joho, J. M. Jose, R. Valenti, and N. Sebe, "Exploiting facial expressions for affective video summarisation," in *Proceedings of the Proceeding of the ACM International Conference*, p. 1, Santorini, Fira, Greece, July 2009.

[33] M. A. Teruel, E. Navarro, V. López-Jaquero, F. Montero, and P. González, "A comprehensive framework for modeling requirements of CSCW systems," *Journal of Software: Evolution and Process*, vol. 29, no. 5, p. e1858, 2017.

[34] F. Montero, V. López-Jaquero, E. Navarro, E. Sánchez, V. López-Jaquero, and E. Sánchez, "Computer-Aided Relearning Activity Patterns for People with Acquired Brain Injury," *Computers & Education*, vol. 57, pp. 1149–1159, 2011.

[35] M. A. Teruel, E. Navarro, and P. González, "Towards an Awareness Interpretation for Physical and Cognitive Rehabilitation Systems," in *Ubiquitous Computing and Ambient Intelligence*, C. R. García, P. Caballero-Gil, M. Burmester, and A. Quesada-Arencibia, Eds., pp. 121–132, Springer International Publishing, 2016.

Hindawi
Mobile Information Systems
Volume 2017, Article ID 7838464, 10 pages
https://doi.org/10.1155/2017/7838464

Research Article

Fuzzy Intelligent System for Patients with Preeclampsia in Wearable Devices

Macarena Espinilla,[1] Javier Medina,[1] Ángel-Luis García-Fernández,[1] Sixto Campaña,[2] and Jorge Londoño[3]

[1] *University of Jaén, Jaén, Spain*
[2] *Universidad Nacional Abierta y a Distancia, Bogotá, Colombia*
[3] *Universidad Pontificia Bolivariana, Medellín, Colombia*

Correspondence should be addressed to Macarena Espinilla; mestevez@ujaen.es

Received 1 June 2017; Revised 4 August 2017; Accepted 10 September 2017; Published 12 October 2017

Academic Editor: Pino Caballero-Gil

Preeclampsia affects from 5% to 14% of all pregnant women and is responsible for about 14% of maternal deaths per year in the world. This paper is focused on the use of a decision analysis tool for the early detection of preeclampsia in women at risk. This tool applies a fuzzy linguistic approach implemented in a wearable device. In order to develop this tool, a real dataset containing data of pregnant women with high risk of preeclampsia from a health center has been analyzed, and a fuzzy linguistic methodology with two main phases is used. Firstly, linguistic transformation is applied to the dataset to increase the interpretability and flexibility in the analysis of preeclampsia. Secondly, knowledge extraction is done by means of inferring rules using decision trees to classify the dataset. The obtained linguistic rules provide understandable monitoring of preeclampsia based on wearable applications and devices. Furthermore, this paper not only introduces the proposed methodology, but also presents a wearable application prototype which applies the rules inferred from the fuzzy decision tree to detect preeclampsia in women at risk. The proposed methodology and the developed wearable application can be easily adapted to other contexts such as diabetes or hypertension.

1. Introduction

We are currently living in the age of data; our everyday lives are surrounded by sensors that capture information associated with objects, humans, or environments such as vision sensors, motion sensors, light sensors, and medical sensors. The so-called Internet of Things (IoT) [1] is all around us, producing a huge amount of data that needs to be automatically organized and processed in order to produce easy to understand reports for the users, as trying to deal with this raw data directly is certainly far from the human capabilities.

Data analysis [2] is at the core of a relevant amount of recently published works, many of them focusing on data mining techniques to extract useful knowledge from the data and to shape it so that users understand what is happening and act accordingly. The research on methods to communicate this knowledge in a user-friendly way has led to the concept of linguistic descriptions [3], which allow humans to abstract useful data into different levels and dimensions, therefore providing interpretability.

In this context, the computing with words (CWW) methodology emulates human cognitive processes to make reasoning processes and decisions in environments of uncertainty and imprecision [4]. The CWW methodology considers that inputs and outputs should be expressed in a linguistic domain in order to be close to human natural language, therefore providing interpretable and understandable results [5, 6].

The application of *soft computing techniques* for the development of methods for data analysis [7] has proved to be very successful. However, these methods produce results that in most cases still require some level of expertise to understand and to use them properly. This is the main motivation for the development of linguistic knowledge extraction techniques, which rely on the principles of fuzzy logic [7] in order

to communicate relevant information obtained from the analysis of the data.

This paper is focused on the health care area, given its relevance nowadays, although the presented techniques could be applied in other areas like wellness [8]. In general, modern health care systems make use of many technological advances to measure biomedical signals that produce large amounts of raw data.

In the context of the early detection of preeclampsia [9, 10], generating user-friendly knowledge to support the diagnosis and monitoring in scalable contexts is a problem that has not been completely solved yet because sometimes a rigid interpretation of the data can lead to misdiagnosis. For example, in [11] it is highlighted that factors like physiology, body size, and variations on the instrumentation can produce differences in the measured heart-rate values that might lead to misdiagnosis.

Therefore, we propose a methodology based on the linguistic approach to extract knowledge by applying a decision tree analysis [12] on a supervised preeclampsia dataset with the aim of early detection of preeclampsia. The proposed methodology is composed of two phases: (i) a linguistic transformation of the dataset to increase the interpretability and flexibility in the analysis of preeclampsia and (ii) the extraction of linguistic knowledge by means of inferring rules using decision trees to classify the dataset.

The use of white-box classification in the cases of decision trees and fuzzy logic [13] provides a wide range of benefits related to real deployments in Health Systems, among which we highlight the following:

(i) *Increased human interpretability.* It is to identify risky conditions and extract interpretable linguistic knowledge for the medical staff, thanks to the intuitive representation of concepts instead of using numeric values [14].

(ii) *Independence from health measuring instruments*. In our approach, the values are translated to degrees of membership in different fuzzy sets, and therefore it is more robust to variations in precision or granularity. We highlight that the flexibility achieved by the fuzzification of the attributes is key for translating the classification from the original dataset based on measurements with traditional health instruments to different contexts and advanced medical devices, such as wearable devices.

(iii) *Inclusion of human observations in the system*. The need for diagnosing preeclampsia in developing countries, where some areas are remote or isolated from health centers, involves integrating human measurements described by itinerant health staff without advanced instruments [15]. In our study case, the age or weight, and even urine infection, can be described by human observation of the patient through linguistic transformation of the data in the decision tree classifier.

The rest of the paper is structured as follows. Section 2 describes the preeclampsia disease and provides some significant data about its incidence worldwide. Section 3 reviews some notions related to mobile information systems and their application to health. Section 4 describes the proposed fuzzy linguistic methodology. Section 5 presents the results of the case study. Section 6 presents a prototype of wearable application which applies the rules inferred from the obtained fuzzy decision tree. Finally, Section 7 summarizes our conclusions and future work.

2. Preeclampsia Disease

In this section, the medical context of preeclampsia is reviewed and global quantitative data is provided in order to highlight the seriousness of the problem.

Preeclampsia (PE) [9, 10] is a multisystemic disorder that occurs during pregnancy, characterized by hypertension and proteinuria (excess of proteins in the urine). PE usually appears after 20 weeks of pregnancy. This condition is one of the most serious complications of pregnancy and also one of the most feared. In severe cases, it endangers the life of both the mother and the fetus [16]. Therefore, it is necessary to diagnose it and begin the treatment as soon as possible. In [17] we also find that severe preeclampsia is associated with an increment of maternal mortality (0.2%) and higher rate of maternal complications (5%) such as convulsions, pulmonary edema, acute renal or liver failures, liver hemorrhage, disseminated intravascular coagulopathies, and strokes. These complications are usually detected in women who develop preeclampsia before the 32nd week of gestation and in those with preexisting medical conditions [18]. PE may be mild or severe, according to the following clinical parameters [19]:

(i) *Mild PE*: Blood pressure of at least 140/90 mmHg on two occasions six hours apart after the 20th week of pregnancy. Proteinuria greater than 300 mg in 24 hours. Moderate edema and urinary volume in 24 hours greater than 500 ml.

(ii) *Severe PE*: Blood pressure greater than 160/90 mmHg on two occasions six hours apart after the 20th week of gestation. Systolic blood pressure greater than 60 mmHg over baseline and diastolic blood pressure greater than 30 mmHg over baseline. Proteinuria greater than 5 g in 24 hours, massive edema, and systemic symptoms such as pulmonary edema, headache, visual disturbances, pain in the right hypochondrium, high level of liver enzymes, or thrombocytopenia.

It is important to emphasize that "blood pressure" is one of the most important factors to be controlled in pregnant women at risk of PE. Therefore, it would be very useful to have a device for monitoring this factor, preferably in real time, and including the possibility of remote monitoring the reports generated by the device.

On the other hand, maternal mortality is a phenomenon that continues occurring despite the efforts of health institutions all over the world. In order to have a general view of the problem, we present here a brief description of the statistics

Table 1: Statistics related to PE in the world in the last 10 years.

Case study	Year	Results
The Woman with Severe Preeclampsia Who Died from Postpartum Complications [23].	2016/2017	PE is responsible for about 14% of maternal deaths per year in the world.
The World Bank. Trends in Maternal Mortality: 1990 to 2010 [24]	2015	The maternal mortality ratio (MMR), defined as the number of maternal deaths per 100000 live births, was estimated at 216 globally.
Competing Risks Model in Screening for Preeclampsia by Maternal Factors and Biomarkers at 30–34 Weeks' Gestation [25]	2013/2014	PE affects 2-3% of all pregnancies and is a major cause of maternal and perinatal morbidity and mortality.
Reducing Maternal Mortality from Preeclampsia and Eclampsia in Low-Resource Countries—What Should Work? [26]	2012	PE is thought to account for about 15% of maternal mortality worldwide, between 20% and 25% of fetal mortality, and with increased risk of asphyxia and preterm delivery for as much as 25% of neonatal mortality.
Review: Biochemical Markers to Predict Preeclampsia [27]	2010/2011	Preeclampsia is one of the most serious pregnancy complications. The worldwide prevalence of PE ranges from 3% to 8% of pregnancies, affecting a total of 8.5 million women worldwide. PE is responsible for about 18% of maternal deaths and up to 40% of fetal mortality.
World Health Statistics 2016: Monitoring Health for the SDGs [28].	2003/2009	The primary causes of maternal deaths are hemorrhage, preeclampsia, eclampsia, sepsis, or infections and indirect causes.

Table 2: PE statistics in Colombia.

Case study	Year	Results
Research Opportunities in Preeclampsia from the Perspective of Primary Prevention: A Reflection [29].	2006	23.8% of maternal mortality cases recorded in health centers were due to this pathology.
Preeclampsia: Facing a Complex Problem from its Physiology [30].	2010	The indicator of maternal mortality was 25.7% for countries such as Colombia and 16.1% for developed countries. Statistics also showed that the main cause of death due to maternal complications in Colombia in 2010 was PE.
Clinical Practice Guidelines for Approaching Pregnancy Associated Hypertensive Complications [31].	2013	35% of maternal mortality was associated with hypertensive disorders such as PE.
Protocol on Public Health Surveillance: Extreme Maternal Morbidity [32].	2014-2015	10,499 cases of maternal death in 2014, increasing by 35% the figures from 2013. By 2015, the number of cases increased to 15,067.

related to maternal mortality in relation to preeclampsia (PE), mainly referred to the last years (see Table 1).

This paper presents a methodology based on the linguistic approach to extract knowledge by applying a decision tree analysis [12] on a supervised preeclampsia dataset from Colombia with the aim of early detection of preeclampsia. For this reason, it is also important to refer to the PE statistics in this country. A statistical summary can be found in Table 2.

It has also been pointed out [20] that the management of PE has not changed significantly over time, possibly as a result of the poor progress in understanding this condition. All these factors motivate our work, which was developed taking as reference a dataset of patients who suffered PE during a five-year time frame, kindly provided by a hospital in Colombia.

In this paper, we analyze this information using data mining techniques and, as a result, we propose a prediction model and a pathology monitoring process that takes into account the values of relevant vital signs, using a fuzzy linguistic system to support the diagnosis of PE. This prediction model has been developed for a mobile device as a prototype of wearable application in order to detect preeclampsia in women at risk.

3. Mobile Applications in Health Care

In this section, we review some notions related to mobile information systems and some applications of mobile systems to health. A mobile application is software designed to run on a mobile device. Typically, mobile applications will access data, devices, and other applications from anywhere [21].

One of the features offered by mobile platforms is real-time connectivity, which has facilitated the creation of various tools that efficiently guide the daily activities of their users. Mobile applications have become the preferred way for users to connect from their devices [22] and have different application areas, such as health, industry, commerce, marketing, entertainment, and sports, among others. For the case that we analyze, the specific domain is the one that refers to the application of mobile information systems to health care. To provide a global idea about some applications developed in this context, Table 3 shows some related works.

We can divide the mobile applications for health purposes into two main groups: applications that are designed for specific purposes (treatment of a single pathology) and general purpose applications that are intended to prevent or attend to emergency situations. The common feature of all

Table 3: Applications of mobile systems to health.

Application	Year	Service
Fetal heart rate monitor using a mobile phone [33]	2009	A mobile application that analyses the fetal heartbeat and calculates the heart rate using a beat-to-beat accuracy algorithm. This data is sent and stored in a server, and then a midwife can examine it through a web browser [34].
Mobile health system for the treatment of pediatric HIV [35]	2009	A Web based EHR system that uses an embedded comprehensive pediatric HIV knowledge base and a clinical decision system. This system allows physicians to integrate clinical information to manage pediatric HIV.
Mobile phone image transmission for diagnosis [36]	2011	A mobile phone with camera functionalities is used as image transmission unit. This system allows communicating with medical image experts in remote locations. The images are sent using the Multimedia Message Service (MMS) via mobile phones.
Mobile phone pulse oximeter [37]	2012	A mobile phone is used to analyze and display the information received from an oximeter placed on a finger. It can aid physicians in detecting clinical events and making decisions [34].
Portable telemedicine unit [38]	2008	A portable telemedicine unit that combines a mobile telemedicine system with a computer server. Communication between these two elements is set via GSM, CDMA, Internet, or satellite. This system is intended for healthcare services in rural or remote regions and can be used for several health services, such as recording, reporting, and teleconsultation [34].

these applications is that all of them use mobile technologies for their operation, including the monitoring through remote transmissions of the pathology or purpose for which they were created.

In this paper, we present a mobile information system in the context of health, especially, a prototype of wearable application to detect preeclampsia in women at risk. Moreover, the most important variables to be monitored are identified; the values of these variables can be collected from health sensors and/or from human observation straightforwardly.

4. Methodology

In this section, a methodology based on a linguistic approach is presented to extract knowledge by applying a decision tree analysis on a supervised preeclampsia dataset. The aim of this process is to allow early detection of preeclampsia.

The dataset contains quantitative data from patients with pregnancy disorders and risk of preeclampsia in Colombia, as well as a human expert diagnosis label for each case.

Firstly, a series of attributes is identified as key factors for diagnosing preeclampsia, based on the experience of health experts and other academic studies. This process led us to the following selection of attributes:

(i) *Age*: age is between 13 and 46 years old, as stated in [39].

(ii) *Body mass index* (ICM): body mass index is between 22 and 38 Kg/m^2, as explained in [40].

(iii) *Trimester of pregnancy*: pregnancy is to be in the second or third trimester.

(iv) *Blood pressure*: diastolic (DBP) or systolic (SBP) between 80 and 200 mmHg is highly related to preeclampsia, according to [41].

(v) *Family history*: if the mother of the patient suffered from preeclampsia, it is labeled as first degree; if only one of the patient's grandmothers suffered from preeclampsia, it is labeled as second degree; otherwise, no label is set. This is pointed out as a risk factor in [42].

(vi) *Socioeconomic stratum*: it is related to preeclampsia because of its consequences on the supplementation of multivitamins and folic acid [43]. It is described with discrete values in the range [1, 4]. The value of 1 is related to the highest socioeconomic stratum and 4 the lowest, which is the most critical for preeclampsia.

(vii) *Race/ethnicity*: according to [44], African-American women are more prone to suffer from severe preeclampsia. This attribute is defined with discrete values (*indigenous*, *African-American*, and *mestizo*).

(viii) *New mother*: this attribute reveals whether it is the first time the patient is pregnant.

(ix) *Proteinuria*: the presence of excess proteins in the urine. It can be measured by a simple dipstick test that returns a value in a scale ranging from 0 to 8 mg/dl. It is also related to preeclampsia diagnosis, as explained in [45].

(x) *Preeclampsia label*: this is the target attribute to classify.

The last attribute, *preeclampsia label*, is defined as a discrete attribute with three possible values: *nonpreeclampsia*, *moderate preeclampsia*, and *severe preeclampsia*.

We highlight the fact that the cases included in the dataset were collected because of their complexity and relevance to preeclampsia detection. No trivial episodes have been included in the dataset, and the diagnosis of preeclampsia is divided into moderate and severe. The complexity of the cases in the dataset should be taken into account before presenting the results.

In the next subsections, the two phases of the proposed methodology are presented.

4.1. Extracting Knowledge by Means of Decision Trees. The first phase of our methodology is focused on extracting

Table 4: Data for trapezoidal shaped representation.

Attribute	Term	Trapezoidal membership function
Systolic pressure	Normal	T (120 mmHg, 120 mmHg, 135 mmHg, 140 mmHg)
Systolic pressure	High	T (135 mmHg, 140 mmHg, 155 mmHg, 160 mmHg)
Systolic pressure	Severe	T (140 mmHg, 150 mmHg, 160 mmHg, 170 mmHg)
Systolic pressure	Very severe	T (160 mmHg, 170 mmHg, 200 mmHg, 200 mmHg)
Diastolic pressure	Normal	T (85 mmHg, 85 mmHg, 85 mmHg, 90 mmHg)
Diastolic pressure	High	T (85 mmHg, 90 mmHg, 105 mmHg, 110 mmHg)
Diastolic pressure	Severe	T (100 mmHg, 105 mmHg, 115 mmHg, 120 mmHg)
Diastolic pressure	Very severe	T (115 mmHg, 120 mmHg, 200 mmHg, 200 mmHg)
Proteinuria	Normal	T (0 g/dl, 100 m/dl, 900 g/dl, 1000 g/dl)
Proteinuria	High	T (500 g/dl, 1000 g/dl, 4500 g/dl, 5000 g/dl)
Proteinuria	Severe	T (2000 g/dl, 4500 g/dl, 5000 g/dl, 5000 g/dl)
Age	Girl	T (13 y, 13 y, 15 y, 18 y)
Age	Young	T (15 y, 20 y, 30 y, 35 y)
Age	Elder	T (30 y, 38 y, 38 y, 38 y)
Age	Risk	T (35 y, 38 y, 40 y, 40 y)
Weight (ICM)	Normal	T (20 Kg/m^2, 20 Kg/m^2, 20 Kg/m^2, 25 Kg/m^2)
Weight (ICM)	Overweight	T (23 Kg/m^2, 25 Kg/m^2, 30 Kg/m^2, 30 Kg/m^2)
Weight (ICM)	Obesity	T (25 Kg/m^2, 30 Kg/m^2, 35 Kg/m^2, 35 Kg/m^2)
Weight (ICM)	Severe obesity	T (30 Kg/m^2, 35 Kg/m^2, 35 Kg/m^2, 35 Kg/m^2)

knowledge from the dataset on pregnant women with high risk of preeclampsia using decision trees. In this stage, decision trees are used to infer rules to classify the dataset.

Among all the standard decision tree strategies, we propose using the C4.5 statistical classifier [46] to analyze the dataset. This classifier is shaped as a tree-like graph for continuous and discrete attributes [47] in which the most relevant attributes are located in upper layers. The hierarchy is built taking into account the entropy of each attribute [48], which is a measurement of the homogeneity and the relevance of the information of an attribute with respect to the target attribute to be classified.

The result of applying a decision tree is a white-box model [49], in which the tree-like graph can be easily understood and modified by humans. Moreover, it can be translated into an inference rule system in which each path, from the root to each end node, represents an induction rule [50].

When analyzing real-world datasets, the model tends to increase the tree size to obtain the best accuracy in the classification. However, this may result in the overfitting of the tree, as well as making it more difficult to interpret; therefore, pruning techniques are applied to reduce the tree complexity [51]. Pruning decreases the size of the tree-like graph by removing those nodes whose instances present low entropy. In Section 5, the effects of applying two different pruning techniques on real data are discussed.

4.2. Linguistic Fuzzy Transformation. The fuzzy linguistic approach is included in the second phase of the classification of the dataset in order to increase the interpretability of the resulting decision tree. Fuzzy logic has been proposed for this stage because of the increased expressiveness of learning methods based on fuzzy logic predicates [52].

We have integrated a fuzzification step in which we relate each numeric attribute with an attribute described using linguistic terms as presented in [13]. In this step, each attribute with continuous values has been translated to a linguistic variable with different terms. The fuzzification consists of describing each attribute by one variable whose terms are related to a membership degree between [0, 1]. For example, *systolic pressure: 138 mmHg* would be described as *systolic pressure is normal (0.8)* and *systolic pressure is high (0.2)*. Next, the continuous value of each attribute is replaced by the linguistic term with the highest membership degree.

The membership functions of the linguistic terms have been provided by health experts, which defined the values and ranges for each attribute based on their experience and the attribute relevance as a risk factor. The attributes, linguistic terms, and membership functions proposed are detailed in Table 4 using a trapezoidal shaped representation, as shown in Figure 1.

5. Evaluation of Results

In this section, a real evaluation of the methodology proposed in this paper is carried out in order to show the efficiency of our proposal.

We have used a dataset from patients with a diagnosis of possible preeclampsia collected in the Departmental Hospital of Nariño (Colombia); this meant an arduous task of compilation with private and public funding. According to medical experts, the dataset contains a representative sample of preeclampsia cases in the last 5 years in this area and consists of 729 records with the attributes and values described in Section 4. We present the results of analyzing

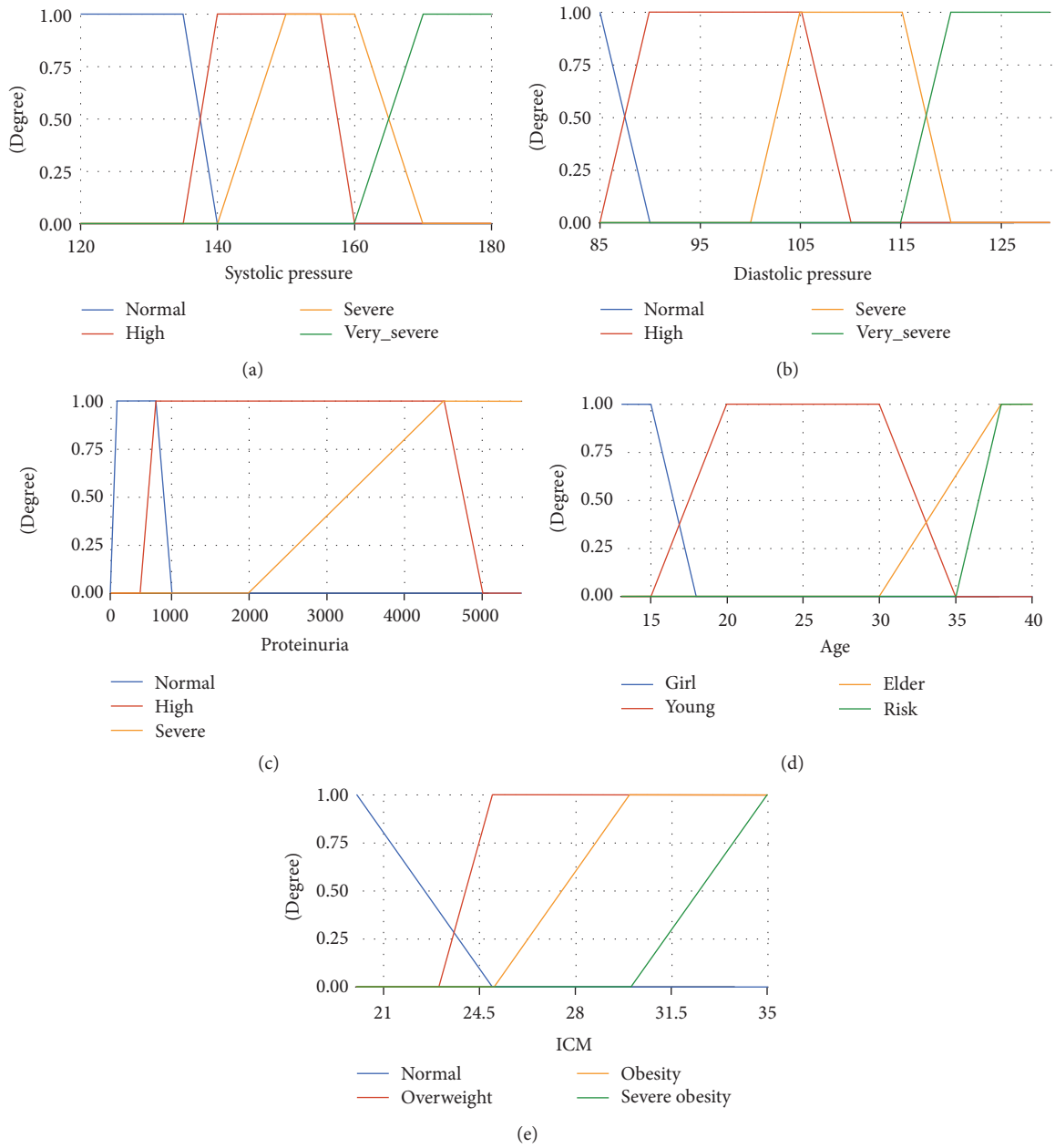


Figure 1: (a) Linguistic terms and membership functions related to systolic pressure. (b) Linguistic terms and membership functions related to diastolic pressure. (c) Linguistic terms and membership functions related to proteinuria. (d) Linguistic terms and membership functions related to the age of the patient. (e) Linguistic terms and membership functions related to ICM.

the dataset using the open source implementation of C4.5 available in Weka [53], called J48 [54].

The initial approach was to classify the overall dataset using simple C4.5 without pruning techniques and without a limit in the number of required instances for creating a tree node. The accuracy obtained was 99.45% but at the cost of a tree size of 598 nodes. As we stated before, the problems of overfitting and interpretability can be reduced at the expense of accuracy using pruning. For this reason, in the second and third attempt, the accuracy ratio is computed on the test set by means of a leave-one-out cross-validation, which is a particular case of f-cross-validation when f is equal to 1. The main advantage of this validation is that all the activities in the dataset are used for training and testing, therefore avoiding the problem of considering how the dataset is divided.

In the second attempt, we applied two pruning techniques (one was applied on the finished tree and the other one was applied during the tree construction); these were the following: (i) reducing the generated nodes with a confidence factor of entropy equal to 0.25 and (ii) disabling nodes with just one instance in construction. This approach produced a notable accuracy of 82.16%, with a tree size of 236 nodes.

Age ≤ 20
  Etnia = mestizo
    Trimester of pregnancy = 3T
      Body Mass Index ≤ 28 THEN Risk of preeclampsia is severe
      Body Mass Index > 28 THEN Risk of preeclampsia is moderate
    Trimester of pregnancy = 2T
      Age ≤ 18 THEN Risk of preeclampsia is moderate
      Age > 18 THEN Risk of preeclampsia is severe

Box 1: Example of rules generated from original data.

patient is a girl
  patient with severe obesity
    Proteinuria is not normal
      patient without antecedents
        Etnia = african-american THEN Risk of preeclampsia is severe
        Etnia = indigenous THEN Risk of preeclampsia is moderate
        Etnia = mestizo THEN Risk of preeclampsia is moderate
      patiente with antecendets THEN Risk of preeclampsia is moderate

Box 2: Example of interpretable rules generated by linguistic approach.

Table 5: Results of the methods proposed for classifying preeclampsia patients.

C4.5 features	Accuracy	Relative size	Advantages	Weak points
(i) No pruning (ii) No limit in number of nodes	0.9945	100%	(i) Good accuracy	(i) Low interpretability because of the tree size (ii) Overfitting (iii) Difficult extrapolation out of the dataset
(i) Postpruning (confidence factor = 0.25) (ii) Online pruning (number of instances/node ≥ 2)	0.8216	39%	(i) Encouraging accuracy (ii) Increased interpretability due to the reduction in the tree size	(i) Difficult extrapolation out of the dataset (ii) Dependence on measuring instruments (iii) Not good accuracy
(i) Fuzzy linguistic representation (ii) Postpruning (confidence factor = 0.25) (iii) Online pruning (number of instances/node ≥ 2)	0.7503	32%	(i) Increased interpretability due to the linguistic approach (ii) Flexibility on measuring instruments (iii) Human observation of attributes	(i) Loss of precision

For the third attempt, we evaluated the fuzzy approach of tree classification, integrating the linguistic terms and membership functions proposed by health experts as described in Section 4.2. It resulted in a tree with 197 nodes and an improvement of interpretability at the expense of a lower accuracy of 75.03%. Obviously, the generation of more scalable classification system based on knowledge involves a loss of accuracy, which is related in [55, 56] as *a reasonable small loss of accuracy for the sake of interpretability*. In exchange, it can be adapted to other contexts using human observation or different health devices, such as a wearable.

We include a fragment of the rules generated from the original data and the linguistic approach, respectively, in Boxes 1 and 2. Note the interpretability of the linguistic approach which can deal with human perception of age and symptoms.

We have summarized the results of analyzing the preeclampsia dataset in Table 5.

6. Development Using Wearable Devices

In this section, we present a prototype of wearable application which contains the rules from the fuzzy decision tree detailed in Sections 4 and 5. It has been developed using the Android Wear platform. As we have described in this paper, the aim of the methodology is to generate medical knowledge that can be applied with the aid of wearable devices in situations in which access to health centers is restricted.

Firstly, the wearable application collects the input values for the attributes to monitor from the patient. This information can be collected in different ways:

(a) (b) (c)

Figure 2: (a) Introducing input information with linguistic terms, (b) getting the measurement value from a wireless health device, including the fuzzification and representation of the value with a pie circle chart (green for *normal* and yellow for *severe*), and (c) representation of the degree of preeclampsia with a pie circle chart (green for *nonpreeclampsia*, yellow for *moderate preeclampsia*, and red for *severe preeclampsia*).

(i) Measurements from wireless health devices, such as wireless blood pressure monitors connected through Bluetooth Low Energy (BLE). In this case, the continuous value of the device is replaced by a linguistic term under the fuzzification described in Section 4.2.

(ii) Linguistic terms related to human interpretation, for example, overweight or obesity.

(iii) Linguistic terms related to human observation of colored strips from portable urine tests. It is successful, for example, for visually evaluating the excess of proteinuria [57] in a cheap and portable way.

Secondly, the linguistic terms are evaluated in the decision tree integrated in the wearable device, providing a matching with the inferred rules.

Finally, the degree of risk of preeclampsia in the unit interval is provided as the degree of membership to each target class (*nonpreeclampsia*, *moderate preeclampsia*, and *severe preeclampsia*). Figure 2 shows the wearable application.

Two features have prevailed in the software development for the wearable device: on the one hand, experimentability, that allows the tool to be used to evaluate and experiment with different linguistic rules generated by different datasets and on the other hand, maintainability, due to the use of frameworks and standard, well-documented programming languages. Due to these features, the software can be easily modified and fixed for further development.

7. Conclusions and Future Work

Early detection of preeclampsia is an important worldwide problem that should be addressed. This paper has presented a solution to support the diagnosis and monitoring of this disease, overcoming the limitation of a rigid interpretation of the data that can usually lead to misdiagnosis. In this paper, biomedical signals are used to identify risk conditions and extract linguistic knowledge for the medical staff. A methodology based on a fuzzy linguistic approach that provides interpretable results is proposed. The proposed methodology is composed of two phases that include a knowledge extraction by means of decision trees and a linguistic transformation of the data. Therefore, this methodology allows a linguistic monitoring in real time. A real evaluation of the proposed methodology has been carried out, providing good results and offering interpretable rules for monitoring. Furthermore, a prototype of a wearable application, which applies the rules from the fuzzy decision tree derived from the analysis, has been presented. Finally, our future work is focused on trying to obtain a preeclampsia dataset from another health center, preferably from another country, and applying the proposed methodology in order to compare and analyze the inferred linguistic rules for the new dataset.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement no. 734355, the Spanish government with the TIN2015-66524-P Project, and the Asociación Universitaria Iberoamericana de Postgrado.

References

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): a vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[2] H. Tanaka, "Fuzzy data analysis by possibilistic linear models," *Fuzzy Sets and Systems. An International Journal in Information Science and Engineering*, vol. 24, no. 3, pp. 363–375, 1987.

[3] A. Ramos-Soto, A. Bugarin, S. Barro, and J. Taboada, "Automatic Generation of Textual Short-Term Weather Forecasts on Real Prediction Data," *Fuzzy Systems, IEEE Transactions*, vol. 23, no. 1, pp. 44–57, 2013.

[4] L. A. Zadeh, "From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions," *IEEE Transactions on Circuits and Systems. I. Fundamental Theory and Applications*, vol. 46, no. 1, pp. 105–119, 1999.

[5] L. Martínez, D. Ruan, and F. Herrera, "Computing with words in decision support systems: An overview on models and applications," *International Journal of Computational Intelligence Systems*, vol. 3, no. 4, pp. 382–395, 2010.

[6] J. M. Mendel, L. A. Zadeh, E. Trillas et al., "What computing with words means to me: Discussion forum," *IEEE Computational Intelligence Magazine*, vol. 5, no. 1, pp. 20–26, 2010.

[7] L. A. Zadeh, "Fuzzy logic, Neural networks, and soft computing," *Communications of the ACM*, vol. 37, no. 3, pp. 77–84, 1994.

[8] T. M. Uribe, "El autocuidado y su papel en la promoción de la salud," *Investigación y Educación en Enfermera*, vol. 17, no. 2, 2013.

[9] C. W. G. Redman, G. P. Sacks, and I. L. Sargent, "Preeclampsia: an excessive maternal inflammatory response to pregnancy," *American Journal of Obstetrics & Gynecology*, vol. 180, no. 2, pp. 499–506, 1999.

[10] L. M. Yamaleyeva and S. H. Lindsey, "Potential for miRNAs as Biomarkers and Therapeutic Targets in Preeclampsia," *Hypertension*, vol. 69, no. 4, pp. 580-581, 2017.

[11] J. M. Goldman, "The challenge of acquiring accurate, complete, near-patient clinical data for data science analysis," in *Data Science Symposium*, p. 43, 2014.

[12] J. R. Quinlan, "Simplifying decision trees," *International Journal of Man-Machine Studies*, vol. 27, no. 3, pp. 221–234, 1987.

[13] M. E. Cintra, M. C. Monard, and H. A. Camargo, "A fuzzy decision tree algorithm based on c4. 5," *Mathware & Soft Computing*, vol. 20, pp. 56–62, 2013.

[14] Y. Zheng, L. Peng, L. Lei, and Y. Junjie, "R-C4.5 decision tree model and its applications to health care dataset," in *Proceedings of the 2005 International Conference on Services Systems and Services Management, ICSSSM'05*, vol. 2, pp. 1099–1103, Chongquing, China, June 2005.

[15] P. Singh, "Employing UNICEF Open Source Software Tools in mHealth Projects in Nicaragua," in *Ubiquitous Computing and Ambient Intelligence*, vol. 10069 of *Lecture Notes in Computer Science*, pp. 286–293, Springer International Publishing, 2016.

[16] C. I. Arroyo Vasquez, "Factores de riesgo independientes para la presencia de preeclampsia," 2014.

[17] B. M. Sibai, "Diagnosis and management of gestational hypertension and preeclampsia," *Obstetrics & Gynecology*, vol. 102, no. 1, pp. 181–192, 2003.

[18] B. M. Sibai, M. A. Villar, and B. C. Mabie, "Acute renal failure in hypertensive disorders of pregnancy. Pregnancy outcome and remote prognosis in thirty-one consecutive cases," *American Journal of Obstetrics and Gynecology*, vol. 162, no. 3, pp. 777–783, 1990.

[19] L. A. V. Egan and S. P. C. Peña, "Conceptos actuales sobre la preeclampsiaeclampsia," *Revista de la Facultad de Medicina UNAM*, vol. 50, no. 2, 2007.

[20] F. A. English, L. C. Kenny, and F. P. McCarthy, "Risk factors and effective management of preeclampsia," *Integrated Blood Pressure Control*, vol. 8, pp. 7–12, 2015.

[21] J. G. Enriquez and S. I. Casas, "Usabilidad en aplicaciones moviles," *Informes Cientificos-Tecnicos UNPA*, vol. 5, no. 2, pp. 25–47, 2014.

[22] L. Ramirez, E. Guillen, and Y. Cifuentes, "Validation strategy to mobile health applications Estrategia de validacion para aplicaciones moviles de salud," *Actas de Ingenieria*, vol. 2, pp. 325–333, 2016.

[23] A. M. Gashi, "The woman with severe preeclampsia who died from postpartum complications," *Revista Cubana de Obstetricia y Ginecologia*, vol. 42, no. 4, pp. 519–523, 2016.

[24] WHO, UNICEF, and UNFPA, *The World Bank. Trends in Maternal Mortality: 1990 to 2010*, WHO, UNICEF, UNFPA, and The World Bank Estimates, 2012.

[25] A. Tsiakkas, Y. Saiid, A. Wright, D. Wright, and K. H. Nicolaides, "Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 30-34 weeks' gestation," *American Journal of Obstetrics and Gynecology*, vol. 215, no. 1, pp. 87.e1–87.e17, 2016.

[26] R. L. Goldenberg, B. Jones, J. B. Griffin et al., "Reducing maternal mortality from preeclampsia and eclampsia in low-resource countries - What should work?" *Acta Obstetricia et Gynecologica Scandinavica*, vol. 94, no. 2, pp. 148–155, 2015.

[27] U. D. Anderson, M. G. Olsson, K. H. Kristensen, B. Åkerström, and S. R. Hansson, "Review: biochemical markers to predict preeclampsia," *Placenta*, vol. 33, pp. S42–S47, 2012.

[28] World Health Organization, *World Health Statistics 2016: Monitoring Health for the SDGs Sustainable Development Goals*, World Health Organization, 2016.

[29] L. A. Díaz Martínez and N. C. Serrano Díaz, "Oportunidades de investigación en preeclampsia, desde la perspectiva de prevención primaria. Un artículo de reflexión. A paper aimed at inviting reflection on," *Revista Colombiana de Obstetricia y Ginecología*, vol. 59, no. 3, pp. 206–215, 2008.

[30] C. Piedrahita Ochoa and B. Agudelo Jaramillo, "Preeclampsia: un problema complejo para enfrentar desde su fisiología," *Medicina & Laboratorio*, vol. 16, no. 11/12, 2010.

[31] A. Cinets, "Guía práctica clínica para el abordaje de las complicaciones hipertensivas asociadas al embarazo. Clinical practice guidelines for approaching pregnancyassociated hypertensive complications," *Revista Colombiana de Obstetricia y Ginecología*, vol. 64, no. 3, pp. 289–326, 2013.

[32] M. L. Ospina, M. E. Martínez, O. E. Pacheco, and H. Q. Bonilla, "Protocolo de Vigilancia en Salud Pública," *Morbilidad Materna Extrema*, 2016.

[33] C. S. Lee, M. Masek, C. P. Lam, and K. T. Tan, "Advances in fetal heart rate monitoring using smart phones," in *Proceedings of the 9th International Symposium on Communications and Information Technology (ISCIT '09)*, pp. 735–740, IEEE Press, Icheon, Republic of Korea, September 2009.

[34] B. M. C. Silva, J. J. P. C. Rodrigues, I. de la Torre Díez, M. López-Coronado, and K. Saleem, "Mobile-health: a review of current state in 2015," *Journal of Biomedical Informatics*, vol. 56, pp. 265–272, 2015.

[35] S. Paul, S. Bhattacharya, A. Sudar et al., "A web-based electronic health care system for the treatment of pediatric HIV," in *Proceedings of the 11th International Conference on e-Health Networking, Applications and Services (Healthcom '09)*, pp. 175–180, Sydney, Australia, December 2009.

[36] L. Bellina and E. Missoni, "M-learning: Mobile phones' appropriateness and potential for the training of laboratory technicians in limited-resource settings," *Health and Technology*, vol. 1, no. 2-4, pp. 93–97, 2011.

[37] D. Dunsmuir, C. Petersen, W. Karlen, J. Lim, G. A. Dumont, and J. M. Ansermino, "The phone oximeter for mobile spot-check," *Anesthesia and Analgesia*, vol. 115, no. 2S, p. 21, 2012.

[38] E. Sutjiredjeki, S. Soegijoko, T. L. R. Mengko, and S. Tjondronegoro, "Development of mobile telemedicine system with multi communication links to reduce Maternal Mortality Rate," in *Proceedings of the 6th IASTED International Conference on Biomedical Engineering, BioMED 2008*, pp. 402–407, ACTA Press, Anaheim, Calif, USA, February 2008.

[39] J. S. Moldenhauer, J. Stanek, C. Warshak, J. Khoury, and B. Sibai, "The frequency and severity of placental findings in women with preeclampsia are gestational age dependent," *The American Journal of Obstetrics and Gynecology*, vol. 189, no. 4, pp. 1173–1177, 2003.

[40] T. E. O'Brien, J. G. Ray, and W.-S. Chan, "Maternal body mass index and the risk of preeclampsia: A systematic overview," *Epidemiology*, vol. 14, no. 3, pp. 368–374, 2003.

[41] M. A. Brown, M. D. Lindheimer, M. de Swiet, A. van Assche, and J.M. Moutquin, "The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP)," *Hypertension in Pregnancy*, vol. 20, no. 1, pp. 9–14, 2001.

[42] R. Arngrimsson, S. Björnsson, R. T. Geirsson, H. Björnsson, J. J. Walker, and G. Snaedal, "Genetic and familial predisposition to eclampsia and pre-eclampsia in a defined population," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 97, no. 9, pp. 762–769, 1990.

[43] S. W. Wen, X. K. Chen, M. Rodger et al., "Folic acid supplementation in early second trimester and the risk of preeclampsia," *American Journal of Obstetrics and Gynecology*, vol. 198, no. 1, pp. 45–e1, 2008.

[44] A. A. Goodwin and B. M. Mercer, "Does maternal race or ethnicity affect the expression of severe preeclampsia?" *American Journal of Obstetrics and Gynecology*, vol. 193, no. 3, pp. 973–978, 2005.

[45] E. Schiff, S. A. Friedman, L. Kao, and B. M. Sibai, "The importance of urinary protein excretion during conservative management of severe preeclampsia," *American Journal of Obstetrics and Gynecology*, vol. 175, no. 5, pp. 1313–1316, 1996.

[46] J. R. Quinlan, *C4. 5: Programs for Machine Learning*, Elsevier, 2014.

[47] J. R. Quinlan, "Improved use of continuous attributes in C4.5," *Journal of Artificial Intelligence Research*, vol. 4, pp. 77–90, 1996.

[48] C. E. Brodley and M. A. Friedl, "Decision tree classification of land cover from remotely sensed data," *Remote Sensing of Environment*, vol. 61, no. 3, pp. 399–409, 1997.

[49] B. Delibasic, M. Vukicevic, and M. I. L. O. Jovanovic, "White-box decision tree algorithms: A pilot study on perceived usefulness, perceived ease of use, and perceived understanding," *International Journal of Engineering Education*, vol. 29, no. 3, pp. 674–687, 2013.

[50] J. R. Quinlan, "Generating Production Rules from Decision Trees," in *IJCAI*, vol. 87, pp. 304–307, 1987.

[51] J. Mingers, "An Empirical Comparison of Pruning Methods for Decision Tree Induction," *Machine Learning*, vol. 4, no. 2, pp. 227–243, 1989.

[52] O. Cordón, "A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems," *International Journal of Approximate Reasoning*, vol. 52, no. 6, pp. 894–913, 2011.

[53] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.

[54] N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria, "Decision tree analysis on j48 algorithm for data mining," *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 6, 2013.

[55] S. Guillaume and B. Charnomordic, "Generating an interpretable family of fuzzy partitions from data," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 3, pp. 324–335, 2004.

[56] J. M. Alonso and L. Magdalena, "Special Issue on Interpretable Fuzzy Systems," *Information Sciences*, vol. 181, no. 20, pp. 4331–4339, 2011.

[57] D. A. De Silva, A. C. Halstead, A.-M. Côté, Y. Sabr, P. von Dadelszen, and L. A. Magee, "Urinary Dipstick Proteinuria Testing: Does Automated Strip Analysis Offer an Advantage Over Visual Testing?" *Journal of Obstetrics and Gynaecology Canada*, vol. 36, no. 7, pp. 605–612, 2014.