

Research Article

Emotion-Aware Assistive System for Humanistic Care Based on the Orange Computing Concept

Jhing-Fa Wang, Bo-Wei Chen, Wei-Kang Fan, and Chih-Hung Li

Department of Electrical Engineering, National Cheng Kung University, Tainan 70101, Taiwan

Correspondence should be addressed to Jhing-Fa Wang, wangjf@mail.ncku.edu.tw

Received 10 February 2012; Accepted 12 April 2012

Academic Editor: Qiangfu Zhao

Copyright © 2012 Jhing-Fa Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mental care has become crucial with the rapid growth of economy and technology. However, recent movements, such as green technologies, place more emphasis on environmental issues than on mental care. Therefore, this study presents an emerging technology called orange computing for mental care applications. Orange computing refers to health, happiness, and physiopsychological care computing, which focuses on designing algorithms and systems for enhancing body and mind balance. The representative color of orange computing originates from a harmonic fusion of passion, love, happiness, and warmth. A case study on a human-machine interactive and assistive system for emotion care was conducted in this study to demonstrate the concept of orange computing. The system can detect emotional states of users by analyzing their facial expressions, emotional speech, and laughter in a ubiquitous environment. In addition, the system can provide corresponding feedback to users according to the results. Experimental results show that the system can achieve an accurate audiovisual recognition rate of 81.8% on average, thereby demonstrating the feasibility of the system. Compared with traditional questionnaire-based approaches, the proposed system can offer real-time analysis of emotional status more efficiently.

1. Introduction

During the past 200 years, the industrial revolution has caused a considerable effect on human lifestyles [1, 2]. A number of changes occurred [3] with the rapid growth of the economy and technology, including the information revolution [3], the second industrial revolution [4], and biotechnology development. Although such evolution was considerably beneficial to humans, it has caused a number of problems, such as capitalism, utilitarianism, poverty gap, global warming, and an aging population [1, 2]. Because of recent changes, a number of people recognized these crises and appealed for effective solutions [5], for example, the green movement [6], which successfully creates awareness of environmental protection and leads to the development of green technology or green computing. However, the green movement does not concentrate on body and mind balance. Therefore, a solution that is feasible for shortening the discrepancy between technology and humanity is of utmost concern.

In 1972, the King of Bhutan proposed a new concept that used gross national happiness (GNH) [7] to describe the standard of living of a country, instead of using gross domestic product (GDP). The GNH has attracted considerable attention because it measured the mental health of people. Similar ideas were also proposed in other works. For example, Andrew Oswald advocated Happiness Economics [8] by combining economics with other research fields, such as psychology and sociology. Moreover, a book entitled “Well-Being” [9], which was written by Daniel Kahneman (a Nobel Prize winner in Economic Sciences in 2002) explained the fundamentals of happy psychology. The common objective of those theories is to upgrade the living quality of humans and to bring more happiness into our daily lives. Recently, the IEEE launched the humanitarian technology challenge (HTC) project (<http://www.ieeehtc.org/>) [10] by sponsoring resource-constrained areas to build reliable electricity and medical facilities. Such an action also highlights the importance of humanistic care. Similar to the HTC project, Intel has supported a center for aging services technologies

(CAST) (<http://www.agingtech.org/>), and its objective is to accelerate development of innovative healthcare technologies. Several academic institutes responded to the trend and subsequently initiated medical care research, such as the “CodeBlue” project at Harvard University [11] and “Computers in the Human Interaction Loop” (CHIL) at Carnegie Mellon University [12]. Inspired by those related concepts [1, 2, 6, 8–12], this study devised a research project for studying the new interdisciplinary “Orange Technology” to promote health, happiness, and humanistic care.

Instead of emphasizing the relations between environments and humans, as proposed by green technology, the objective of the orange computing project is to bring more care or happiness to humans and to promote mental wellness for the well-being of society.

Orange computing is an interdisciplinary field that includes computer science, electrical engineering, biomedical engineering, psychology, physiology, cognitive science, and social science. The research scope of orange computing contains the following.

- (1) Health and security care for the elderly, children, and infants.
- (2) Care and disaster relief for people in disaster-stricken areas.
- (3) Care for low-income families.
- (4) Body-mind care for people with physiological and psychological problems.
- (5) Happiness indicator measurement and happiness enhancement.

To demonstrate the concept of orange computing, a case study on a human-machine interactive and assistive system for emotion care was investigated in this study. The proposed system is capable of recognizing human emotions by analyzing facial expressions and speech. When the detected emotion status exceeds a threshold, an alarm will be sent to a doctor or a nurse for further diagnosis and treatment.

The remainder of this paper is organized as follows: Section 2 introduces the orange computing models; Section 3 presents a discussion of a case study on the emotion recognition system for care services; Section 4 summarizes the performance of the proposed method and the analysis results; lastly, Section 5 offers conclusions.

2. Related Work and Orange Computing Concept

Orange computing originates from health informatics, and it contains two research topics: one is physiological care and the other psychological care. Both of the two topics focus on enhancing humans’ physical and mental health, enriching positive emotions and finally bring more happiness to others [13, 14]. The physiological and psychological care models of orange computing are similar to the health model in medical expert systems [15, 16], which have been well developed and commonly used in health informatics over several decades.

In a medical expert system, when a user inputs a query through the interface, the system can automatically search predefined knowledge databases and consult with relevant experts or doctors. After querying databases or merging opinions of experts, the system subsequently replies to the user with an appropriate response. In traditional medical expert systems, database querying and feedback usually involve semantic understanding techniques and delicate interface design [17–19], so that users do not feel inconvenient during the process. However, in some telemedical care systems, such as [20], knowledge databases and feedback mechanisms are replaced with caregivers for better interactivity. Recently, expert systems have gradually integrated knowledge-based information management systems with pervasive computing [21]. Although such systems have been prototyped and modeled in several studies [22, 23], they have not been deployed. However, the abovementioned ideas have spurred the development of orange computing.

Happiness informatics, or the happiness model, is the key characteristic of orange computing. Similar to the health model, the happiness model also requires a user input and a predefined database. The input is commonly measured from the biosignals or behavior of a user, for example, facial expressions, emotional speech, laughter, body gestures, gaits, blood pressure, heartbeat rates, electroencephalograms (EEGs), electrocardiograms (ECGs), and electromyograms (EMGs) [24, 25]. With such information, the happiness model can help users evaluate their emotional status in various applications. Nevertheless, it is quite challenging to determine the manner in which to combine those data and determine emotional status [26–28].

3. Case Study

This section demonstrates a technological application for daily humanistic care in home environments. The system uses contactless multimodal recognition techniques to measure positive emotion degree of users. The recognition results can be logged into the database and sent to analysts for further processing. As shown in Figure 1, the ambient devices of the proposed system include multiple audiovisual sensors, a service robot, and a smart TV. The robot is a self-propelled machine with four wheels and serves as a remote agent between users and the server. To interact with users, it is equipped with audiovisual sensors, loudspeakers, and a touch screen. Similar to the robot, the TV is also used for interacting with users.

After the ambient sensors receive signals from users, data are subsequently sent to a processing server through a cloud network. The workflow of the data processing procedures comprises three stages, as follows: the first and second stages are the audiovisual recognition, and the last stage is the feedback stage. The detail of each stage is described as follows.

3.1. Visual Recognition. At the image processing stage, as shown in Figure 2, after video streams are captured by the camera, Haar-like features [29] are extracted and sent to

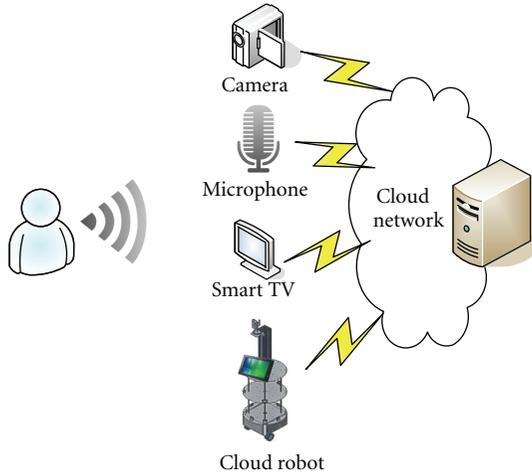


FIGURE 1: Framework of the system.

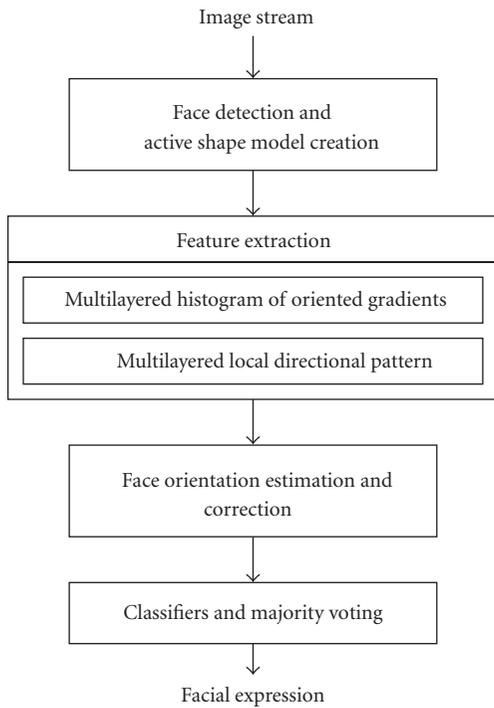


FIGURE 2: Workflow of the image processing stage.

AdaBoost classifiers [29] to detect user faces. Subsequently, the system uses the Active Shape Model, which was proposed by Cootes et al. [30], to model facial regions. Thus, facial regions can be represented by a set of points using the point distribution model.

A novel feature called “Multilayered Histogram of Oriented Gradients” (MLHOGs) is proposed in this study to generate reliable characteristics for estimating facial expressions. The MLHOGs are derived from Histograms of Oriented Gradients (HOGs) [31] and Pyramid Histograms of Oriented Gradients (PHOGs) [32]. Let $f(x, y)$ represent the pixel of coordinate x and y , G denote gradients, W refer

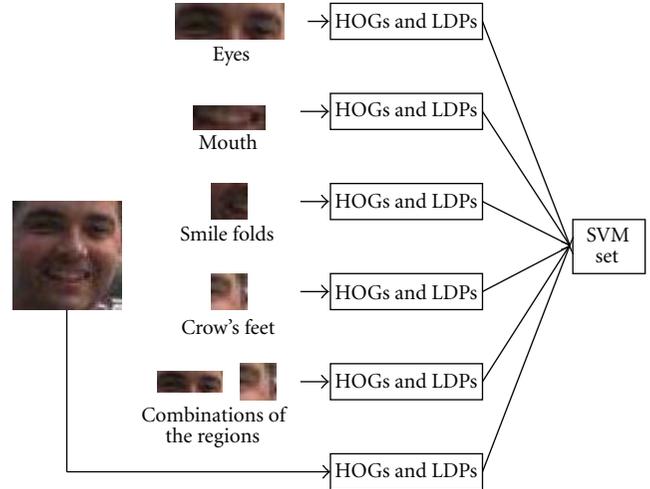


FIGURE 3: Concept of multilayered histogram of oriented gradients and multilayered local directional patterns (the facial image is extracted from the MPLab GENKI database).

to the weight of a coordinate, and θ be edge directions. The histogram of oriented gradients can be expressed as follows:

$$\begin{aligned}
 G_{x,y}^{\text{Horizontal}} &= 2f(x+1, y) - 2f(x-1, y) + f(x+1, y+1) \\
 &\quad - f(x-1, y+1) + f(x+1, y-1) \\
 &\quad - f(x-1, y-1), \\
 G_{x,y}^{\text{Vertical}} &= 2f(x, y+1) - 2f(x, y-1) + f(x+1, y+1) \\
 &\quad - f(x+1, y-1) + f(x-1, y+1) \\
 &\quad - f(x-1, y-1), \\
 W_{x,y} &= \left((G_{x,y}^{\text{Horizontal}})^2 + (G_{x,y}^{\text{Vertical}})^2 \right)^{1/2}, \\
 \theta_{x,y} &= \tan^{-1} \left(\frac{G_{x,y}^{\text{Horizontal}}}{G_{x,y}^{\text{Vertical}}} \right).
 \end{aligned}
 \tag{1}$$

After gradients are computed, a histogram of edge directions is subsequently created to collect the number of pixels that belongs to a direction.

Unlike pyramid histogram of oriented gradients, which concentrates on fixed rectangular shapes inside an image, the proposed MLHOGs are modeled by object-based regions of interest (ROIs), such as eyes, mouths, noses, and combinations of ROIs. Furthermore, each object-based ROI has a dedicated classifier for recognizing the same type of ROIs. A concept example of multilayered histogram of oriented gradients and multilayered local directional patterns is illustrated in Figure 3.

Similar to the proposed MLHOGs, our study also develops a new texture descriptor called “Multilayered Local Directional Pattern” for enhancing recognition rates. Such multilayered directional patterns are computed according to

“edge responses” of pixels, which are based on the same concept of Jabit’s feature, “Local Directional Patterns (LDPs)” [33]. The difference is that the proposed method focuses on patterns at various ROI levels. Computation of multilayered local directional patterns is listed as follows:

$$\mathbf{R}_\psi = \mathbf{F} * \mathbf{M}_\psi, \quad (2)$$

$$\varepsilon_\psi = \sum_{\forall \text{block}_{3 \times 3}} \text{LDP}_{\text{Binary Code}}(\mathbf{R}_\psi), \quad (3)$$

where \mathbf{F} is the input image, \mathbf{M} means eight-directional Kirsch edge masks like Sobel operators, \mathbf{R} stands for edge responses of \mathbf{F} , ψ represents eight directions, and ε is the number of edge responses in a designated direction. Before the system accumulates the edge responses of \mathbf{R} using (3), an LDP binary operation [33] is imposed on \mathbf{R} to generate an invariant code. A one-by-eight histogram is adopted to collect the edge responses in the eight directions. In the proposed multilayered local directional patterns, only edge responses in objects of interest are collected, so that the histogram differs from ROIs to ROIs.

In addition to upright and full frontal faces, this work also supports roll/yaw angle estimation and correction. The active shape model can label facial regions. Relative positions, proportions of facial regions, and orientations of nonfrontal faces can be measured properly with the use of spatial geometry. Once the direction is determined, corresponding transformation matrices are applied to the nonfrontal faces for pose correction.

At the end of the image processing stage, multiple Support Vector Machines (SVMs) are used to classify facial expressions. Each of the SVMs is trained to recognize a specific facial region. The classification result is generated by majority voting.

3.2. Audio Recognition. Audio signals and visual data have a considerable effect on deciphering human emotions. Therefore, the audio processing stage focuses on detecting emotional speech and laughter to extract emotional cues from acoustic signals. The workflow at this stage is illustrated in Figure 4.

First, silence segments in audio streams are removed by using voice activity detection (VAD) algorithm. Subsequently, an autocorrelation method called “Average Magnitude Difference Function” (AMDF) [34] is used to extract phoneme information from acoustic data. The AMDF can effectively estimate periodical signals, which are the main characteristics of speech, laughter, and other vowel-based nonspeech sounds. The AMDF is derived as follows:

$$\begin{aligned} \tau^* &= \arg \min_{\tau} \text{AMDF}(\tau), \\ \text{AMDF}(\tau) &= \sum_{t=0}^{T-t-1} |S(t) - S(t + \tau)|, \end{aligned} \quad (4)$$

where S represents one of the segments in the acoustic signal, T is the length of S , t denotes the time index, and τ is the shifting length. After $\text{AMDF}(\tau)$ reaches the minimum, a

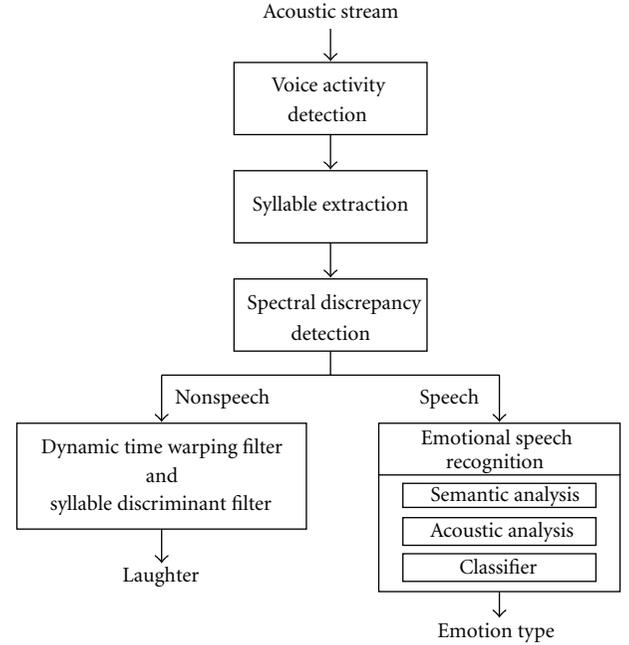


FIGURE 4: Workflow of the audio processing stage.

phoneme $P = [S(\tau^*), S(\tau^* + 1), \dots, S(2\tau^*)]$ can be acquired by extracting indices from $S(\tau^*)$ to $S(2\tau^*)$.

Algorithm 1 expresses the process of syllable extraction when phonemes of a signal are determined.

In the next step, to classify signals into their respective categories, energy and frequency changes are used as the first criteria to separate speech from vowel-based nonspeech because spectral discrepancy of speech is relatively smaller in most cases.

Compared with other vowel-based nonspeech, the temporal pattern of laughter usually exhibits repetitiveness. To detect such patterns, this study uses cascade filters, which consist of a Dynamic Time Warping (DTW) filter [35] and a syllable discriminant filter, to compute similarities of the input data. With the use of Mel-frequency cepstral coefficients (MFCCs), the Dynamic Time Warping filter can find out desired signals by matching them with the samples in the database. The signals that successfully pass through the first filter are subsequently input to the second filter. The syllable discriminant filter compares each input sequence with predefined patterns by using the inner product operation. When the score of an input is higher than a threshold, the input is labeled as laughter.

For emotional speech recognition, this study follows previous works [36–38] and extracts prosodic and timbre features from speech to recognize emotional information in voices. Tables 1 and 2 show the acoustic features used in this system.

In addition to the acoustic features, this study also uses the keyword spotting technique to detect predefined keywords in speech because textual data offer more emotion clues than acoustic data. After detecting predefined keywords in utterances, the system iteratively computes the association

```

Initialization
  designating the beginning phonemes  $P_{start}$ ;
For each phoneme  $P_m$ 
begin
  If Similarity ( $P_{start}, P_m$ ) <  $\delta_{Similarity}$ 
  If Distance ( $P_{m-1}, P_m$ ) >  $\delta_{Distance}$ 
     $m$  is the end of asyllable;
end

```

ALGORITHM 1: Algorithm for syllable extraction.

TABLE 1: Timbre features.

Type	Parameter
1st–3rd formants	Frequency
	Mean
	Standard deviation
	Medium
Spectrum related	Bandwidth
	Centroid
	Spread
	Flatness

TABLE 2: Prosodic features.

Type	Parameter
Pitch and energy related	Maximum value
	Minimum value
	Mean
	Medium
	Standard deviation
	Range
Duration related	Coefficients of the linear regression
	Speech rate
	Ratio between voiced and unvoiced regions
	Duration of the longest voiced speech

degree between the detected keyword and each emotion category.

Let i represent the index of the emotion categories; j denote the index of the detected keyword in the sentence corpus; ω_j refer to the detected keyword; $\Gamma(\omega_j, c_i)$ represent the occurrence of ω_j in category c_i ; $\Gamma(\omega_j)$ denote the number of sentences containing (ω_j).

The association degree can be defined as

$$e_i(\omega_j) = \frac{\Gamma(\omega_j, c_i)}{\Gamma(\omega_j)} \times \frac{\sum_i \Gamma(\omega_j, c_i)^2}{\Gamma(\omega_j)^2}, \quad (5)$$

where the first part of the equation is the weighting score, and the second part is the confidence score of ω_j (see [39] for detailed information). The textual feature vector is

subsequently combined with the acoustic feature vector and sent into a classifier (AdaBoost) for training and recognition.

3.3. Feedback Mechanism. After completion of the audiovisual recognition stage, the system generates three results along with their classification scores. One of the three results is the detected facial expression, another is the detected vocal emotion type, and the other is laughter. The classification scores are linearly combined with the recognition rates of the corresponding classifiers and finally output to users. Additionally, the recognition result is logged in the database 24 hours a day. A user can browse the curve of emotion changes by viewing the display. The system is also equipped with a telehealthcare module. Personal emotion status can be sent to family psychologists or psychiatrists for mental care. The service robot can serve as an agent between the cloud system and users, providing a remote interactive interface.

4. Experimental Results

This study conducted an experiment to test audiovisual emotion recognition to assess the performance of our system. Only positive emotions, including smiling faces, laughter, and joyful voices, were tested in the experiment.

At the evaluation of the facial expression stage, 500 facial images containing smiles and nonsmiles were manually selected from the MPLab GENKI database (<http://mplab.ucsd.edu/>). The kernel function of the SVM was the radial basis function, and the penalty constant was empirically set to one. Furthermore, 50% of the dataset was used for training, and 50% was used for testing. During the evaluation of laughter recognition, a database consisting of 84 sound clips was created by recording the utterances of six people. Eighteen samples from these 84 clips were the sound of people laughing. After removing silence parts from all of the clips, the entire dataset was subsequently sent into the system for recognition. For emotional speech recognition, this research used the same database as that in our previous work [39]. The speech containing joyful and nonjoyful emotions was manually chosen and parsed to obtain their literal information and acoustic features. Finally, these features were inputted into an AdaBoost classifier for training and testing.

Figure 5 shows a summary of the experimental results of our system, in which the vertical axis denotes accuracy rates,

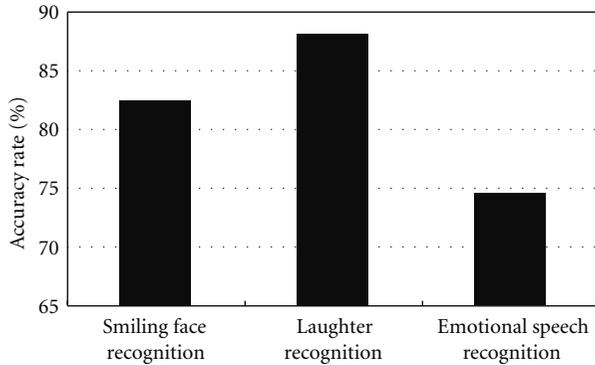


FIGURE 5: Accuracy rates of the audiovisual recognition.

and the horizontal axis represents recognition modules. As shown in the figure, the accuracy rate of smile detection can reach 82.5%. The performance of laughter recognition can also achieve an accuracy rate as high as 88.2%. Compared with smile and laughter recognition, although the result of emotional speech recognition reached 74.6%, such performance is comparable to those of related emotional speech recognition systems. When combined with the test result of emotional speech recognition, the overall accuracy rate can reach an average of 81.8%.

The following experiment tests whether the proposed system can help testees remind and evaluate their emotional health status as caregivers do. During the experiment, total ten persons were selected from the sanatorium and the hospital to test the system for a week. The age of the participants ranges from 40 to 70 years old. The audiovisual sensors were installed in their living space, so that the emotional data can be acquired and analyzed in real time. For privacy, the sensors captured behavior only during 10:00 to 16:00. To avoid generating biased data, each testee was not aware of the locations of the sensors and the testing details of the experiment. Furthermore, after the system analyzed the data, the medical doctors and nurses helped testees complete questionnaires. The questionnaire contained total ten questions, nine of which were irrelevant to this experiment. The remaining question was the key criterion that allowed the testees to give a score (one(unhappy)–five(happy)) to their daily moods.

The questionnaire scores are subsequently compared with the estimated emotional status of the proposed system. To obtain the estimated emotional score, the proposed method firstly calculates the duration of smiling face expressions, joyful speech, and laughter of the testees. Next, a ratio can be computed by converting the duration into a one-to-five rating scale based on the test period.

The correlation test in Figure 6 shows performance of the questionnaire approach and the proposed system. The vertical axis represents the questionnaire result, whereas the score of the proposed system is listed on the horizontal axis. All the samples are collected from the testees. Closely examining the scatterness in this figure reveals that Pearson's correlation coefficient reaches as high as 0.27. This implies that our method is analogous with the questionnaire-based

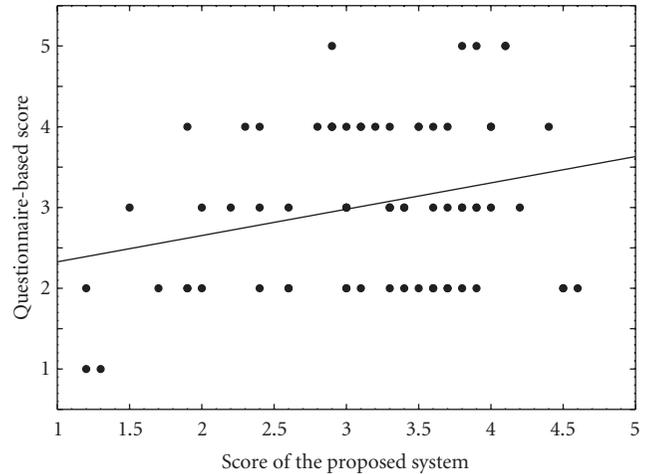


FIGURE 6: Correlation test of the scores between the questionnaire approach and the proposed system. The horizontal axis represents the evaluation result of the proposed system, whereas the vertical axis means the result of the questionnaire. The slope of the regression line is 0.33, and Pearson's correlation coefficient is 0.27.

approach. Moreover, two groups of the scores in the linear regression analysis reflect a linear rate of 0.33. Above findings indicate that the proposed method can allow computers to monitor users' emotional health, subsequently assisting caregivers in reminding users' psychological status and saving more human resources.

5. Conclusion

This paper presents a new concept called orange computing for health, happiness, and humanistic care. To demonstrate the concept, a case study on the audiovisual emotion recognition system for care services is also conducted. The system uses multimodal recognition techniques, including facial expression, laughter, and emotional speech recognition to capture human behavior.

At the facial expression recognition stage, multilayered histograms of oriented gradients and multilayered local directional patterns are proposed to model facial features. To detect patterns of laughing sound, two cascade filters consisting of a Dynamic Time Warping filter and a syllable discriminant filter are used in the acoustic processing phase. Furthermore, when classifying emotional speech, the system combines textual, timbre, and prosodic features to calculate association degree to predefined emotion classes.

Three analyses are conducted for evaluating recognition performance of the proposed methods. Experimental results show that our system can reach an average accuracy rate of 81.8%. Concerning the feedback mechanism, data from the real-life test indicate that our method is comparable to the questionnaire-based approach. Additionally, correlation degree between two methods is as high as 0.27. The above results demonstrate that the proposed system is capable of recognizing users' emotional health and thereby providing an in-time reminder for them.

In summary, orange computing hopes to arouse awareness of the importance of mental wellness (health, happiness, and warming care), subsequently leading more people to join the movement, to share happiness with others, and finally to enhance the well-being of society.

Acknowledgments

This work was supported in part by the National Science Council of the Republic of China under Grant no. 100-2218-E-006-017. The authors would like to thank Yan-You Chen, Yi-Cheng Chen, Wei-Kang Fan, Chih-Hung Li, and Da-Yu Kwan for contributing the experimental data and supporting this research.

References

- [1] M. C. Jensen, "The modern industrial revolution, exit, and the failure of internal control systems," *Journal of Applied Corporate Finance*, vol. 22, no. 1, pp. 43–58, 1993.
- [2] F. Dunachie, "The success of the industrial revolution and the failure of political revolutions: how Britain got lucky," *Historical Notes*, vol. 26, pp. 1–7, 1996.
- [3] Y. Veneris, "Modelling the transition from the industrial to the informational revolution," *Environment & Planning A*, vol. 21, no. 3, pp. 399–416, 1990.
- [4] J. Hull, "The second industrial revolution: the history of a concept," *Storia Della Storiografia*, vol. 36, pp. 81–90, 1999.
- [5] P. Ashworth, "High technology and humanity for intensive care," *Intensive Care Nursing*, vol. 6, no. 3, pp. 150–160, 1990.
- [6] P. Gilk, *Green Politics is Eutopian*, Lutterworth Press, Cambridge, UK, 2009.
- [7] S. B. F. Hargens, "Integral development—taking the middle path towards gross national happiness," *Journal of Bhutan Studies*, vol. 6, pp. 24–87, 2002.
- [8] A. J. Oswald, "Happiness and economic performance," *Economic Journal*, vol. 107, no. 445, pp. 1815–1831, 1997.
- [9] D. Kahneman, E. Diener, and N. Schwarz, *Well-Being : The Foundations of Hedonic Psychology*, Russell Sage Foundation Publications, New York, NY, USA, 1998.
- [10] K. Passino, "World-wide education for the humanitarian technology challenge," *IEEE Technology and Society Magazine*, vol. 29, no. 2, p. 4, 2010.
- [11] K. Lorincz, D. J. Malan, T. R. F. Fulford-Jones et al., "Sensor networks for emergency response: Challenges and opportunities," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 16–23, 2004.
- [12] A. Waibel, "Speech processing in support of human-human communication," in *Proceedings of the 2nd International Symposium on Universal Communication (ISUC '08)*, p. 11, Osaka, Japan, December 2008.
- [13] J.-F. Wang, B.-W. Chen, Y.-Y. Chen, and Y.-C. Chen, "Orange computing: challenges and opportunities for affective signal processing," in *Proceedings of the International Conference on Signal Processing, Communications and Computing*, pp. 1–4, Xian, China, September 2011.
- [14] J.-F. Wang and B.-W. Chen, "Orange computing: challenges and opportunities for awareness science and technology," in *Proceedings of the 3rd International Conference on Awareness Science and Technology*, pp. 538–540, Dalian, China, September 2011.
- [15] Y. Hata, S. Kobashi, and H. Nakajima, "Human health care system of systems," *IEEE Systems Journal*, vol. 3, no. 2, pp. 231–238, 2009.
- [16] K. Siau, "Health care informatics," *IEEE Transactions on Information Technology in Biomedicine*, vol. 7, no. 1, pp. 1–7, 2003.
- [17] K. Kawamura, W. Dodd, and P. Ratanaswasd, "Robotic body-mind integration: Next grand challenge in robotics," in *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN '04)*, pp. 23–28, Kurashiki, Okayama, Japan, September 2004.
- [18] P. Belimpasakis and S. Moloney, "A platform for proving family oriented RESTful services hosted at home," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 690–698, 2009.
- [19] L. S. A. Low, N. C. Maddage, M. Lech, L. B. Sheeber, and N. B. Allen, "Detection of clinical depression in adolescents' speech during family interactions," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 574–586, 2011.
- [20] C. Yu, J.-J. Yang, J.-C. Chen et al., "The development and evaluation of the citizen telehealth care service system: case study in Taipei," in *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '09)*, pp. 6095–6098, Minneapolis, Minn, USA, September 2009.
- [21] J. B. Jørgensen and C. Bossen, "Executable use cases: requirements for a pervasive health care system," *IEEE Software*, vol. 21, no. 2, pp. 34–41, 2004.
- [22] A. Mihailidis, B. Carmichael, and J. Boger, "The use of computer vision in an intelligent environment to support aging-in-place, safety, and independence in the home," *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 3, pp. 238–247, 2004.
- [23] Y. Hata, S. Kobashi, and H. Nakajima, "Human health care system of systems," *IEEE Systems Journal*, vol. 3, no. 2, pp. 231–238, 2009.
- [24] Y. Gizatdinova and V. Surakka, "Feature-based detection of facial landmarks from neutral and expressive facial images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 135–139, 2006.
- [25] R. A. Calvo and S. D'Mello, "Affect detection: an interdisciplinary review of models, methods, and their applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, 2010.
- [26] C. Busso and S. S. Narayanan, "Interrelation between speech and facial gestures in emotional utterances: a single subject study," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 8, pp. 2331–2347, 2007.
- [27] Y. Wang and L. Guan, "Recognizing human emotional state from audiovisual signals," *IEEE Transactions on Multimedia*, vol. 10, no. 4, pp. 659–668, 2008.
- [28] Z. Zeng, J. Tu, B. M. Pianfetti, and T. S. Huang, "Audio-visual affective expression recognition through multistream fused HMM," *IEEE Transactions on Multimedia*, vol. 10, no. 4, pp. 570–577, 2008.
- [29] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1511–1518, Kauai, Hawaii, USA, December 2001.
- [30] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models—their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

- [31] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 886–893, San Diego, Calif, USA, June 2005.
- [32] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," in *Proceedings of the 6th ACM International Conference on Image and Video Retrieval (CIVR '07)*, pp. 401–408, Amsterdam, Netherlands, July 2007.
- [33] T. Jabid, M. H. Kabir, and O. Chae, "Local Directional Pattern (LDP)—a robust image descriptor for object recognition," in *Proceedings of the 7th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS '10)*, pp. 482–487, Boston, Mass, USA, September 2010.
- [34] C. K. Un and S.-C. Yang, "A pitch extraction algorithm based on LPC inverse filtering and AMDF," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 6, pp. 565–572, 1977.
- [35] J.-F. Wang, J.-C. Wang, M.-H. Mo, C.-I. Tu, and S.-C. Lin, "The design of a speech interactivity embedded module and its applications for mobile consumer devices," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 2, pp. 870–876, 2008.
- [36] S. Casale, A. Russo, G. Scebbba, and S. Serrano, "Speech emotion classification using Machine Learning algorithms," in *Proceedings of the 2nd Annual IEEE International Conference on Semantic Computing (ICSC '08)*, pp. 158–165, Santa Clara, Calif, USA, August 2008.
- [37] C. Busso, S. Lee, and S. Narayanan, "Analysis of emotionally salient aspects of fundamental frequency for emotion detection," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 4, pp. 582–596, 2009.
- [38] N. D. Cook, T. X. Fujisawa, and K. Takami, "Evaluation of the affective valence of speech using pitch substructure," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 1, pp. 142–151, 2006.
- [39] Y.-Y. Chen, B.-W. Chen, J.-F. Wang, and Y.-C. Chen, "Emotion aware system based on acoustic and textual features from speech," in *Proceedings of the 2nd International Symposium on Aware Computing (ISAC '10)*, pp. 92–96, Tainan, Taiwan, November 2010.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

