

## Research Article

# Sketch-Based Spatial Queries for the Retrieval of Human Locomotion Patterns in Smart Environments

**Gamhewage C. de Silva, Toshihiko Yamasaki, and Kiyoharu Aizawa**

*Department of Information and Communication Engineering, The University of Tokyo, Tokyo 113-8656, Japan*

Correspondence should be addressed to Gamhewage C. de Silva, chamds@hal.t.u-tokyo.ac.jp

Received 15 March 2009; Accepted 7 July 2009

Recommended by Qi Tian

A system for retrieving video sequences created by tracking humans in a smart environment, by using spatial queries, is presented. Sketches made with a pointing device on the floor layout of the environment are used to form queries corresponding to locomotion patterns. The sketches are analyzed to identify the type of the query. Directional search algorithms based on the minimum distance between points are applied for finding the best matches to the sketch. The results are ranked according to the similarity and presented to the user. The system was developed in two stages. An initial version of the system was implemented and evaluated by conducting a user study. Modifications were made where appropriate, according to the results and the feedback, to make the system more accurate and usable. We present the details of the initial system, the user study and the results, and the modifications thus made. The overall accuracy of retrieval for the initial system was approximately 93%, when tested on a collection of data from a real-life experiment. This is improved to approximately 97% after the modifications. The user interaction strategy and the search algorithms are usable in any environment for automated retrieval of locomotion patterns. The subjects who evaluated the system found it easy to learn and use. Their comments included several prospective applications for the user interaction strategy, providing valuable insight for future directions.

Copyright © 2009 Gamhewage C. de Silva 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.

## 1. Introduction

Multimedia retrieval and summarization for smart environments is an active research area with several applications such as surveillance, study of human behavior, and taking care of the elderly [1]. In most of these applications, it is necessary to track the movement of persons in the environment and retrieve media showing their behavior. Depending on the environment and the application, this can be performed using only the image data, or with the help of other sensors [2, 3]. The results are usually indexed by date, time and person (if recognized) to speed up retrieval.

In some of these applications, such as long term surveillance and monitoring of patients suffering from dementia [4], it is often necessary to search a large collection of tracking data for a particular pattern of movement. Such queries can be categorized as *spatial queries* [5]. This is usually performed by reducing the search space using other criteria and viewing the tracking results manually to retrieve

the desired locomotion pattern. Ability to query a collection of tracking data by a particular locomotion pattern will greatly enhance the efficiency of retrieval, and facilitate identification and analysis of long term behavioral patterns. Therefore, facilitation of spatial querying of human movement patterns in a smart environment solves an important research problem, that is, often faced by researchers working in this area.

However, facilitating spatial queries for a smart environment is a difficult task due to two main problems. First, there should be an intuitive and nonrestricting way to input queries on human locomotion patterns into a computer. Second, algorithms for searching for specific movements have to be designed and developed.

Sketching is a common method used by humans to specify or describe patterns of movement. With several common factors such as area, distance and direction, sketching and locomotion has an intuitive mapping between them. Sketching is a simple activity that almost everybody

is capable of performing. Despite different sketching habits and techniques, many people are capable of interpreting sketches made by others; this suggests that there are some intuitive and widely accepted notations for sketching. All the computer platforms that are widely used today have user interaction capabilities that support sketching. Therefore, sketching is a highly prospective candidate for specifying locomotion patterns and synthesizing queries regarding the same.

In this research, we propose a novel user interaction strategy for spatial querying of locomotion patterns in smart environments. The user specifies a locomotion pattern to be searched for by sketching on a floor plan of the environment. Sketching is performed on a graphical user interface with a pointing device. Different types of queries are designed to enable searching for different locomotion patterns with simple sketches, intuitively and without ambiguity. We also design and implement algorithms for retrieving locomotion patterns represented by these queries.

We implement the proposed spatial querying system in a two-bedroom smart home [6] equipped with a large number of stationary cameras and microphones. Pressure-based sensors mounted on the floor are activated as people move inside the house. At the current status of a research project on multimedia experience retrieval from this environment, it is possible to retrieve personalized video for multiple moving persons by changing cameras and microphones automatically as they move to different regions of the house [7]. The first step of creating such videos is an agglomerative hierarchical clustering algorithm that segments the floor sensor data to a set of sequences that correspond to multiple moving persons inside the house. The algorithm first uses the spatial and temporal distribution of the sensor data to cluster the data into footsteps. These footsteps are then combined to form sequences corresponding to different persons. A video for each footstep sequence is created by automatically selecting cameras and microphones to ensure that the persons are seen and heard throughout the video. The algorithm can accurately segment the movement of multiple persons who are walking or standing on the house floor at the same time. The details of this algorithm and the experimental results are reported in [7].

However, this system facilitates only temporal queries. For example, it is possible to search for video showing people walking inside the house between 6:00 PM and 7:00 PM on a given date. We intend to enhance the capability of the above system by incorporating spatial queries to retrieve video correlated to locomotion patterns. Our objective is to facilitate queries such as: "Retrieve the video clips showing the people who walked from the living room to the study room in the morning of the 1st of May 2007". Using the proposed user interaction strategy, we submit spatial queries by sketching paths on the floor layout of the house. The floor sensor data corresponding to the video clips are then searched for similar paths. The best matches are retrieved and shown to the user as video and footstep sequences. We evaluate the performance of the algorithm using a data set recorded during a real life experiment with actual residents.

The proposed user interaction strategy is designed based on a two-stage process. First, the user strategy is designed and implemented as an initial system that can be used for retrieval of locomotion patterns. We design and conduct user studies to both identify sketch notations used for specifying locomotion patterns, and to evaluate the usability of this system. Based on user feedback, we redesign the user interaction strategy and the search algorithms in order to achieve a better solution.

The rest of this article is organized as follows: Section 2 is a brief review of related work; Section 3 describes the initial system, and the results of quantitative evaluation; Section 4 presents the user study conducted for both requirements acquisition and qualitative evaluation; the results of the user study are presented in Section 5. Section 6 discusses these results, describes the changes made to the system, and presents the results of retrieval using the final system; Section 7 concludes the article with suggestions for future directions.

## 2. Related Work

The research area of smart homes has been a growing area, combining advances in several technologies such as sensing, computing and storage [8]. Current research on smart and ubiquitous environments can be divided into two major categories. One aims at providing services to the people in the environment by detecting and recognizing their actions. Examples include *Aware Home* [9] and other Smart Home projects focusing on assistive environments [10, 11]. The other research category aims at storing and retrieving media at different levels ranging from photos to experiences [7, 12]. There has been a recent growth in research on modeling and visualization of human activities in smart environments, to support both these types of work. Aipperspach et al. [13] use smoothed *n-grams* to model and predict human behavior in a smart environment. Ivanov et al. [14] use *tracklets* for modeling and visualization of human movement. Jaimes et al. attempt to facilitate efficient retrieval of multimedia information using visualization of memory cues [15]. Ability to make spatial queries on people's movement can greatly enhance the functionality of systems such as the above.

There has been some research towards a framework for spatial querying of locomotion patterns. Egenhofer [16] demonstrated how imprecise spatial queries can be dealt with in a comprehensible manner, using topological relations. A relational algebra is proposed for verifying the consistency of the resulting topological representations. Gottfried [4] uses a locomotion base and a set of relations to represent locomotion patterns, with emphasis on healthcare applications. However, an effective user interaction strategy for submitting spatial queries is essential for utilizing the above framework for effective retrieval of locomotion patterns.

So far, there has been very little research on user interfaces for spatial querying. Ivanov and Wren [5] use simple spatial queries for video retrieval from surveillance cameras. However, the functionality of this interface is limited to specifying the direction of movement along a corridor.

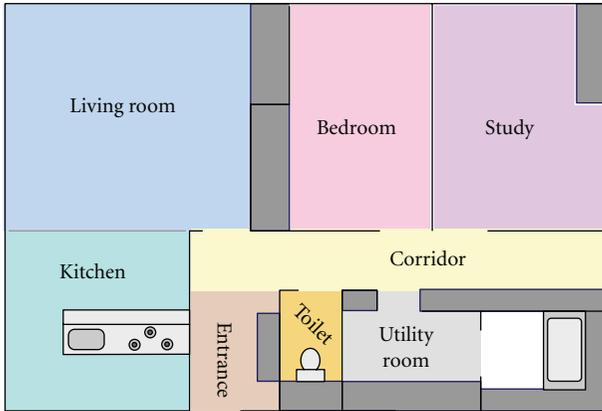


FIGURE 1: Partitioning the smart home floor into regions.

In the following sections of this article, we describe how we approach the problem of spatial querying of locomotion patterns in smart environments based on the framework proposed in previous work, and complement it with a user interaction strategy that facilitates effective searching. We also present user studies designed for gathering information required for designing a user interaction strategy, that is, intuitive and efficient, and evaluating the proposed system.

### 3. Description of System

**3.1. User Interaction Strategy.** Our objective here is to facilitate querying by allowing the users to specify the path they want to search for, by sketching that path on a diagram of the house floor plan. In order to design an interface that is versatile, easy to use and intuitive, we took the following, user-centered approach.

When a person is standing or walking inside a house, he is always referred to as being in a “region”. Therefore, it is desirable to facilitate specification of regions during querying. For this purpose, we partitioned the house floor plan into the regions labeled in Figure 1. We selected the following three types of possible spatial queries to be facilitated by the proposed user interaction strategy:

- (1) *query type 1*: walking/standing within a selected region (e.g., “inside the living room”),
- (2) *query type 2*: walking from one region to another, irrespective of the path taken (e.g., “from the living room to the kitchen”),
- (3) *query type 3*: walking along a specific path.

The user interaction strategy should allow the user to submit all three types of queries in a simple and intuitive manner. There should be no ambiguity between any two types of queries. The complexity of the input is also important. Since queries of types 1 and 2 are less specific, it is desirable for them to be easier to enter.

To query for footstep sequences within a given region, the user scribbles within that region on the floor plan

(Figure 2(a)). Since this is the type of query with the least amount of detail, it is sensible to use a simple gesture for specifying this query.

To query for movement between any two regions, the user simply draws a line between the two of them, without considering the walls that partition the regions (Figure 2(b)). Since movement through such partitioning is impossible, only the starting and ending regions of the lines are useful as meaningful inputs to a query. The user finds the query easy to enter and intuitive, as he/she needs to be concerned about only those two regions. However, it should be noted that there can be some ambiguity when the start and end regions are not well partitioned. This is resolved later during the search algorithm.

When querying for a specific path, the user sketches the path along the house plan (Figure 2(c)). The path can be either within the same region, or across several regions. This query takes more effort from the user, and returns fewer and more specific results.

**3.2. Algorithms for Searching.** The personalized video from the ubiquitous home is created by selecting cameras and microphones automatically for footstep sequences of different persons [17]. The attributes contained in the footstep sequence data are sufficient to fully represent the video clip [18]. Therefore, it is possible to retrieve video by querying only the collection of footstep sequences using the parameters obtained from the inputs, namely, the date, time interval and the sketch.

Each footstep sequence is an ordered set of four-dimensional data elements with the following variables:

- (i)  $X$  coordinate of the position of the sensor,
- (ii)  $Y$  coordinate of the position of the sensor,
- (iii) time stamp indicating when the sensor was activated,
- (iv) duration that the sensor was active.

The  $X$  and  $Y$  coordinates are specified in millimeters, starting from the bottom left corner of the house floor as seen in Figure 1. The proposed search algorithm uses only the coordinates and the time order of the elements for retrieval of these sequences. The duration of the activations apparently has no intuitive mapping with the time taken to input the sketch.

The array of pixel coordinates contained in the sketch has to be preprocessed before searching. First, the points are transformed from pixel coordinates to the house floor coordinates. Ideally, this should be possible using a linear transform if the house floor layout on the interface is drawn to scale. However, due to calibration errors in floor sensor data, minor adjustments are needed. The ordered set of points  $P = \{p_1, p_1, \dots, p_n\}$  is submitted as input to the search algorithms.

The next step is to identify the type of the query by analyzing the distribution of points in  $P$  and the time consumed for making the sketch. Type 1 queries are identified using the distribution of the points. Type 2 queries are distinguished by identifying the crossing of partitions between regions. When

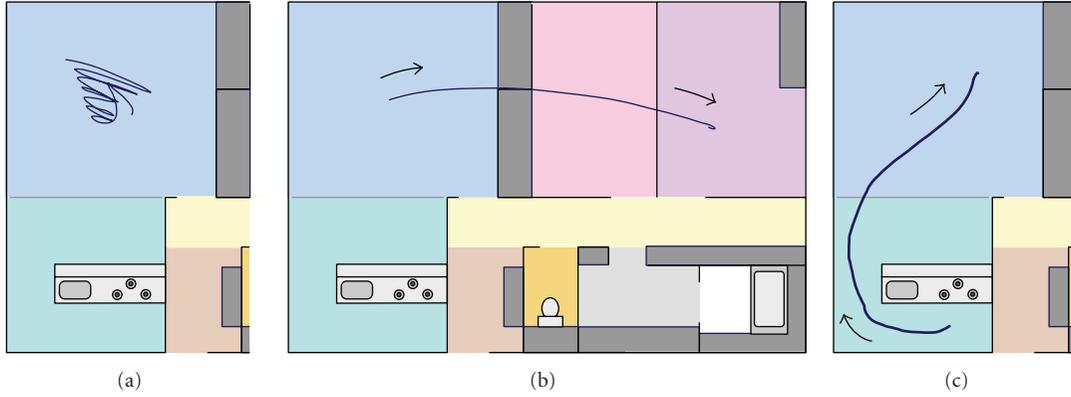


FIGURE 2: Different types of spatial queries.

there is an ambiguity between types 2 and 3, it is resolved using the time consumed to draw the sketch since type 2 queries are faster to sketch.

*Query Type 1.* To perform this query type, we retrieve all footstep sequences present in the specified region, from the collection of tracking data. If a time interval is specified, the results are filtered using that interval. The results are ordered by the starting time of the sequences.

*Query Type 2.* First, we retrieve all the footstep sequences that include floor sensor data from both regions, and filter them by the specified time interval. Thereafter, the following algorithm is applied to each candidate path  $C = \{c_1, c_2, \dots, c_m\}$  thus selected.

- (1) Starting from  $c_1$ , scan  $C$  until the first point in the starting region,  $c_s$ , is found.
- (2) Within the subsequence  $C' = \{c_s, c_{s+1}, \dots, c_m\}$ , find the first point in the ending region,  $c_e$ .
- (3) Select  $C'' = \{c_s, c_{s+1}, \dots, c_e\}$  as a match.

Again, the matches are ordered by the starting time stamp.

*Query Type 3.* We start the search by selecting all the footstep sequences recorded during the time interval that the user specified. For each of the candidate paths  $C = \{c_1, c_2, \dots, c_m\}$ , selected as above, the following algorithm is applied.

- (1) Set overall mean distance  $D = 0$ .
- (2) For the first point  $c_1$  in the selected candidate path  $C$ , find the closest point  $p_a$  in  $P$ .
- (3) Add the Euclidean distance between  $c_1$  and  $p_a$  to  $D$ .
- (4) Repeat the steps 1 and 2 for the next point in  $C$  and  $P' = \{p_a, p_{a+1}, \dots, p_n\}$ , until all points in  $C$  are used in the calculation.
- (5) Divide  $D$  by  $m$ .
- (6) If  $D < T_D$ , select  $C$  as a match.

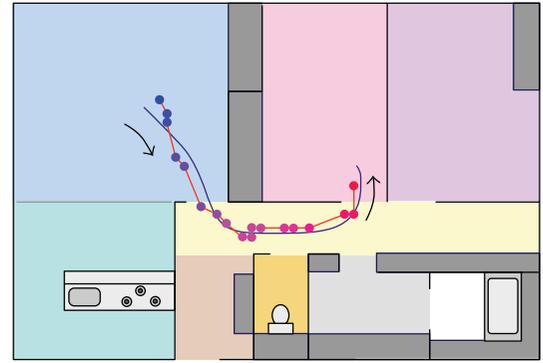


FIGURE 3: A search query of type 3, with the best matching result.

This algorithm looks for paths with less deviation from the sketched path, while preserving direction. The threshold value  $T_D$  is the maximum mean deviation permitted between two paths for them to be matched, and depends on the accuracy of tracking for the environment considered. For this smart home,  $T_D$  was empirically set to 360 mm, equal to twice the resolution of pressure sensors on the floor. The paths with different starting or ending regions to those of the sketch are removed from the set of matches, to prevent false retrievals of much shorter paths with good overlap.

The matched paths are presented to the user in ascending order of the overall mean distance. Figure 3 shows a sketch made by a user (the curved path), and the retrieved path that matched it best (the piecewise linear path). The dots marked along the retrieved path correspond to the locations of the footsteps. The dots change color from blue to red, indicating the direction of the retrieved path.

*3.3. User Interface Design.* We designed a graphical user interface based on the above strategy for retrieval of personalized video from the ubiquitous home. This interface is based on the concept of *hierarchical media segmentation* [7], to facilitate more accurate retrieval based on interactive querying. Figure 4 is a screenshot of the user interface with numbered steps for a retrieval task. The inputs consist of



FIGURE 4: User interaction with system: the circled numbers 1 to 5 correspond to steps 1 to 5 described in text.

a date input, a time line, and a drawing of the house floor layout. All these inputs can be specified using only a pointing device. When the user selects a particular date (step 1), the time line shows one-minute segments where footstep sequences were present during that day. The user then draws a line segment on the time line to specify the time interval, aided by this visualization of the results (step 2). If necessary, the user can now see a summary of all footsteps within the house during this time interval (as seen by the dots on the house floor plan in Figure 4). This helps the user to identify any unusual pattern, or directly see whether the desired pattern might be available for this time slot. Thereafter, the user sketches the path on the house floor plan (step 3). If the user does not select the date, the entire collection of data will be used as input for the search. If a time interval is not specified, the entire day is considered. Upon completing the sketch, the system retrieves the results and shows a summary of results. This summary consists of the set of footsteps drawn on the house floor plan, and the first frame of the corresponding video. The user can browse through this summary (step 4), and view the video clip and footsteps (step 5).

All the pixels along the sketched path and the date and time interval (if submitted) are recorded as inputs for the search algorithm described in the following section. The time to make the sketch is recorded as an additional input.

**3.4. Quantitative Evaluation.** The results retrieved by the first two types of queries are straightforward, since they are direct queries on a relational database followed by a linear search. In order to evaluate the performance of the search algorithm for query type 3, we conducted the following experiment. The house floor was partitioned into *subregions*

with an area of  $90 \times 90$  cm (corresponding to  $5 \times 5$  floor sensors). We decided to evaluate the system by searching for only those paths between a pair of these subregions. Retrieval of sequences with the correct starting and ending subregions was used as an objective measure of the accuracy of retrieval.

The system was tested on a selected set of 94 footstep sequences obtained from 12 hours of data gathered during a “real-life experiment”, where a family of three members stayed in ubiquitous home for 10 days. A set of 56 pairs of subregions were selected by observing these data.

During evaluation, five paths between each selected pair of subregions were sketched and results were retrieved. Both the instances where wrong paths were retrieved and correct paths were missed were recorded. The precision  $P$ , recall  $R$  and balanced  $F$ -measure  $F$  for retrieval were calculated as:

$$\begin{aligned} P &= \frac{N_C}{N_C + N_M}, \\ R &= \frac{N_C}{N_C + N_O}, \\ F &= \frac{2PR}{P + R}, \end{aligned} \quad (1)$$

where  $N_C$  is the number of correctly retrieved video clips,  $N_O$  is the number of clips that were not retrieved, and  $N_M$  is the number of mistakenly retrieved clips.

The precision of retrieval was 92.5% and the recall was 98.8%. The balanced  $F$ -measure was 95.2%. Most of the mistakenly retrieved clips were candidate paths that are shorter than the sketch but match well with the corresponding segment of the sketch.

## 4. User Study

We conducted a user study on sketching spatial queries, with the following objectives.

- (1) Analyze how people sketch locomotion patterns and identify the common sketch types and places where they disagree.
- (2) Find out how people interpret sketches representing locomotion patterns.
- (3) Evaluate the initial system described in Section 3, and identify changes that are necessary to improve its usability.
- (4) Acquire feedback related to both the initial system and sketching locomotion patterns.
- (5) Identify future directions and prospective applications of sketch-based querying of locomotion patterns.

Since it was not possible to find an existing method of evaluation available to fulfill all of the above, we designed and conducted our own user study. This study consisted of five sections. The first three sections consisted of three different tasks related to sketching locomotion patterns. The fourth section was a usability study of the initial system. The fifth section contained questions regarding sketch-based querying, and allowed the subjects to write their comments freely. The following subsections of the article describe the above in detail.

*4.1. Sketching a Locomotion Pattern Based on a Textual Description.* The objective of this section is to identify the notations that people use and the difficulties they encounter, when sketching a locomotion pattern that they haven't experienced or seen. The section consists of 16 sketching tasks. In each task, the test subjects read a textual description of an instance of locomotion in the house (e.g., "walking from the living room to the study") and sketched it on a floor plan of the house. The descriptions were selected in such a way that they describe different locomotion patterns. Some descriptions with a certain degree of ambiguity were deliberately included.

All the sketches were made on answer sheets with preprinted house floor layouts. After the tasks were completed, the subjects were asked whether there were any particular movement patterns that were difficult to sketch. They were allowed to describe freely (in writing), if there were such difficulties or comments.

*4.2. Sketching Locomotion Based on Observation.* The objective of this section is to identify the notations that people use and difficulties they encounter when sketching a locomotion pattern that they have seen or experienced. The section consisted of nine sketching tasks. In each task, the subject watched an animation showing a person moving on the house floor plan. These animations were created by observing patterns of movement exhibited by the residents of the house during previous real-life experiments [7]. Creating

animations was necessary since it is difficult for a nonresident to interpret the movement by merely observing multi-camera video. After watching the animation, the subject interpreted the movement, and sketched it on the house floor layout.

All the sketches were made on answer sheets with preprinted house floor layouts. After the tasks were completed, the subjects were asked whether there were any particular movement patterns that were difficult to sketch. They were allowed to describe freely (in writing), if there were such difficulties or comments.

*4.3. Interpreting the Queries on the Proposed System.* The objective of this section is to examine the users' ability to learn a sketching notation for querying locomotion patterns, and identify ambiguities and difficulties from a user's perspective. While the first two sections consisted of general sketching tasks, this section was based on the initial system for retrieving locomotion patterns. The notation of the queries used in the initial system was explained to the subject, with examples. The subject was allowed to try at least one example query from each type by himself, to familiarize with the system before the actual tasks began.

The section consisted of six tasks. During each task, the subject was shown a screen capture of a spatial query on the system, and asked to interpret the query and its type. While interpreting the type of a query is a task for the system, not the user, this question helps us to understand if there are situations where the user is not sure how to sketch a query, due to either ambiguity or lack of intuitiveness. Six screen captures, consisting of two queries from each type, were shown to the users in random order.

After the tasks were completed, the following questions were asked.

- (1) Were there any movement patterns that were difficult to interpret? Describe briefly.
- (2) Out of the sketch types used in this system, which type do you think is the most useful when specifying activity inside a house?

The first question was asked to identify any ambiguities or counter-intuitive notations in the current set of query types. The second question intends to seek for any user preferences, which if found can be used for putting more effort in to designing such queries.

*4.4. Usability Study of the Proposed System.* After using the system further if they thought it necessary, the subjects rated its usability by answering a questionnaire, based on the guidelines established by Chin et al. [19]. A seven-point response scale was used with 1 being the worst rating (very poor performance) and 7 the best (very good performance).

*4.5. Feedback and Comments.* In this section of the experiment, the subjects answered the following general questions:

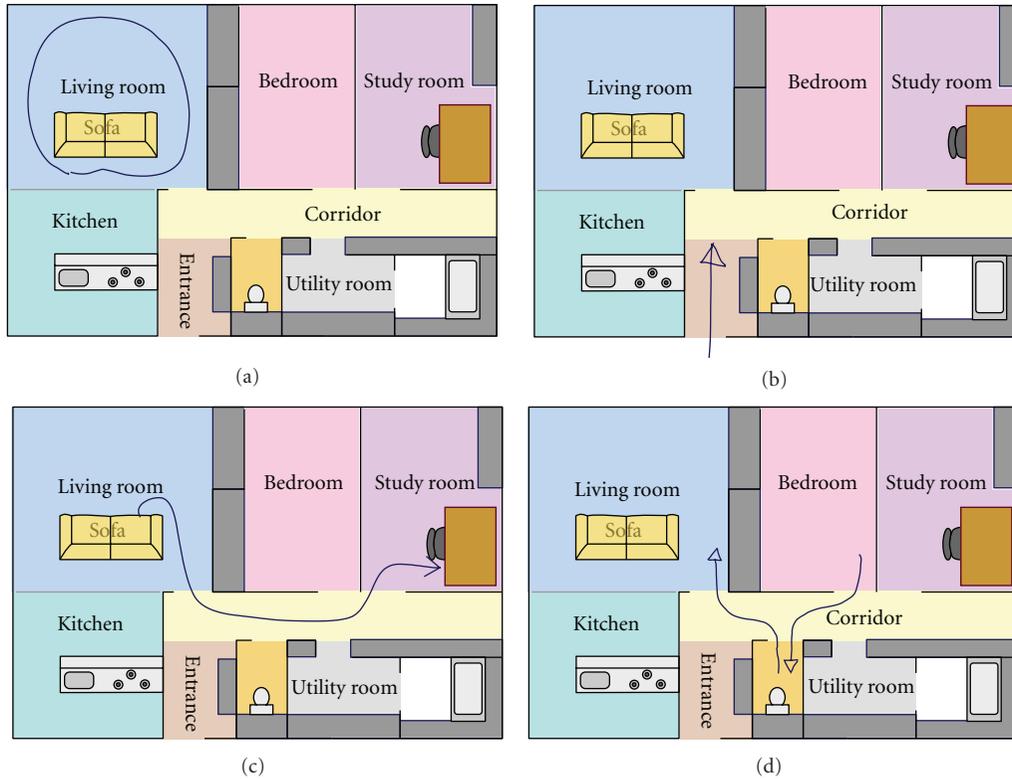


FIGURE 5: Some example sketches by test subjects for different movement patterns: (a) walking inside the living room, (b) entering the house, (c) walking from the sofa to the table in the study, and (d) walking from the bedroom into the toilet and entering the living room after a short pause.

- (1) Do you think it is easier to specify a pattern of movement inside a house by using a sketch than a verbal description?
- (2) What are the other applications that you would suggest for a sketch-based interface that can accept movement patterns as input?

After answering the questions, the subjects were given the opportunity to make additional comments and suggestions in free format.

**4.6. Experimental Procedure.** The procedure of the user study followed the guidelines used in a study for identifying how people sketch geographical information, by Blaser [20]. Sixteen voluntary subjects took part in the study. All of them were regular computer users. However, none of the subjects was involved with the design or using of spatial queries.

Each subject was briefed about the task at the beginning of each section, and also provided with written instructions. One of the authors was available throughout the experiment to provide additional clarifications if needed. Animations and video clips were replayed when the subjects found it necessary. Breaks, if the subjects needed any, were allowed between sections.

The subjects answered all the sections on answer sheets provided to them. In addition to the responses on paper,

the time taken for the experiments and the stroke order of sketches were recorded.

The subjects took 24 to 40 minutes to complete the experiment. The average time consumed was 30.3 minutes. This time included short breaks between sections.

## 5. Results

The responses gathered during the user study consisted of sketches, numerical ratings, and textual descriptions. Figure 5 shows some example sketches made by different subjects (redrawn electronically for clarity). In addition to these direct responses, information regarding the stroke order was also recorded. The following subsections summarize the results of each section of the experiments, and our inferences based on the same.

**5.1. Sketching a Locomotion Pattern Based on a Textual Description.** The textual descriptions provided in this section specified several primitive types of locomotion patterns; entering a region, moving within a specified region/s, moving between regions, and so forth. We attempt to identify sketching notations used by the subjects by separating notations for these primitives. We use the comments by the subjects to identify the difficulties and ambiguities in sketching a locomotion pattern.

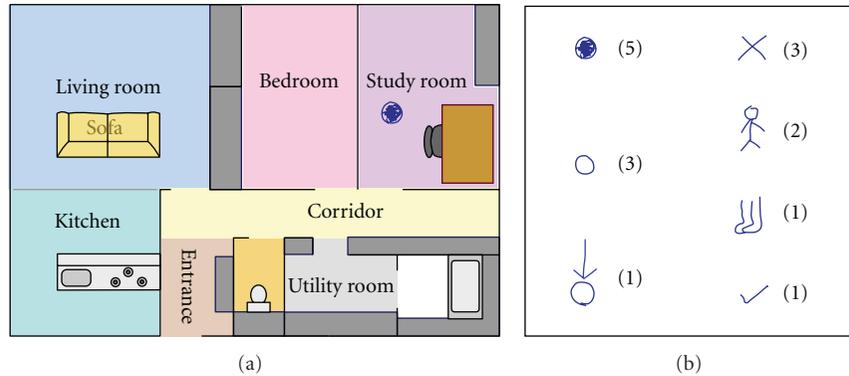


FIGURE 6: Sketch notations for staying at a location.

Figure 6 shows the sketch notations used by subjects to specify the primitive “stationary” (“standing” or “sitting”, as specified in the descriptions of the user study). Figure 6(a) is an example sketch for the description “standing inside the study room”. Figure 6(b) summarizes the notations used by the subjects for this primitive. The numbers shown in parentheses indicate the number of subjects who used each notation.

Most of the subjects (9 out of 16) used a small circle to indicate the location related to the description. A cross was the next most common, and there was no other common notation.

Figure 7 shows the sketch notations used by subjects to specify locomotion within a region. Figure 7(a) is a sketch made by a subject to represent “walking inside the living room”. Figure 7(b) summarizes the responses by all the subjects.

It is evident that most of the subjects (14 of 16) used closed or near-closed curves to indicate movement within a region. Only one subject used a notation similar to that used by the initial system (type 1 query). Two of the subjects used arrowheads to represent movement.

Figure 8 summarizes the sketch notations used by the subjects to indicate movement between two regions. Figure 8(a) is an example sketch, and Figure 8(b) summarizes the notations in the same format as Figures 6(b) and 7(b). The small, numbered arrows in Figure 9(b) show the direction and the order of the strokes used in sketching.

Most of the subjects (12 of 16) used arrowheads to indicate the direction of movement, while the others implied the direction by the direction they used when sketching the line. The subjects drew arrowheads in different styles and stroke orders, as seen in Figure 9.

For queries involving more than two regions (e.g., “Walking from the bedroom to the toilet, and then to the living room”), 9 out of 16 subjects used two line segments indicating a break at the toilet, whereas the others used only one line segment. While sketching these queries, 11 of the 16 subjects used arrowheads to indicate direction.

Entering a region was indicated using an arrow or a line leading into the region, while leaving the region was sketched

using an arrow or a line leading away from a region. Again, 12 of 16 subjects indicated the direction using an arrowhead.

The subjects stated that they encountered difficulties in sketching the following descriptions (the number of subjects providing each response is shown in parentheses):

- (i) entering corridor (8),
- (ii) passing through the corridor in either direction (6),
- (iii) standing inside the study room (4),
- (iv) from living room to kitchen (2).

The main reason for a description to be found “difficult to sketch” was the lack of specific information required for sketching. For instance, in the first description listed above, the users wished to know which entrance was used to enter the corridor. There are several entrances to the corridor, making it difficult to sketch. For the second, some of them wished to know both the entrance and the exit. For the third, the location where the person is standing is not given. Most of the users marked the circle toward the center of the room. However, it is evident that the subjects were less concerned about incomplete information once there was sufficient detail regarding the main action performed. Although the last description above does not mention the exact locations in the rooms, the main action described was “walking (moving)”, and only two subjects found it difficult to sketch.

**5.2. Sketching Locomotion Based on Observation.** The users found it much easier to complete this section, as demonstrated by their comments. There was no ambiguity in interpretation, since the subjects were able to see the movement by themselves. The notations used by most of the subjects were consistent with what they used during Section 1. There were five instances with differences in sketch notation. However, these belonged to different users and different queries. Having no correlation, the differences were treated as negligible.

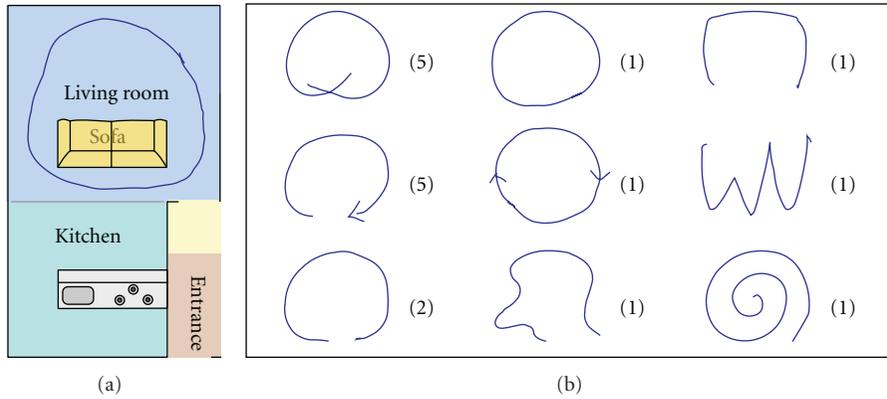


FIGURE 7: Sketch notations for locomotion within a region.

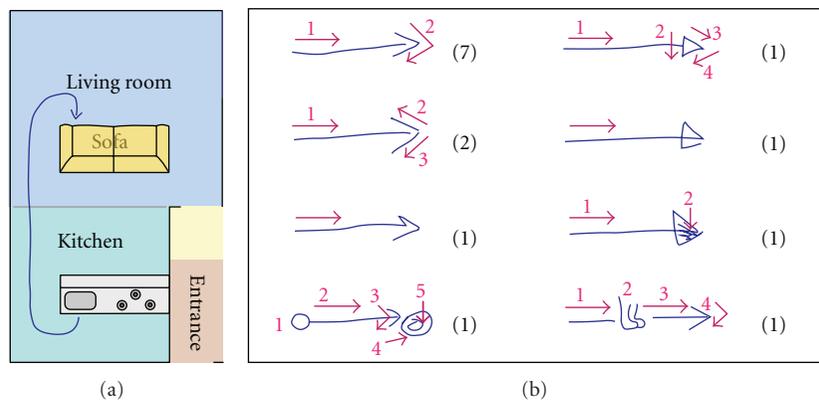


FIGURE 8: Sketch notations for movement between regions using arrows.

Five of the subjects desired to show pauses of the moving person in their sketches. Figure 9 shows an example sketch with a pause (Figure 9(a)) and the notations used by those subjects for showing a pause (Figure 9(b)). It is evident that the majority of the subjects (11 out of 16) ignored the pauses when describing locomotion patterns.

5.3. *Interpreting the Queries on the Proposed System.* The users were able to make an accurate interpretation of the locomotion pattern specified by the sketch in most cases. The number of errors within the set of 96 subject-queries (16 users × 6 queries) was only 6. The result confirms the fact that people find it easy to interpret a sketch once an intuitive notation is agreed upon.

Asked whether there were any patterns that were difficult to interpret, the subjects reported problems in identifying the particular query type in one of the queries (Figure 10). In this case, the lack of a partitioning wall between the living room and the kitchen made it difficult to identify whether the query meant any path from the living room to the kitchen, or the specific path as drawn in the sketch. Most of the users selected the correct answer by considering the speed that the sketch was drawn.

The subjects disagreed heavily on which type of querying is useful for specifying locomotion patterns inside a house. The following are the responses, with the number of respondents in parentheses:

- (i) type 1 (3),
- (ii) type 2 (4),
- (iii) type 3 (5)
- (iv) all types (4).

Some of the subjects preferred the ability to specify minute details and hence preferred type 3. Some others were more interested in “region level locomotion information” provided by query type 2.

Providing additional comments, two of the subjects mentioned that query type 1 (scribbling in a region, corresponding to the presence within the region) is difficult both to sketch and interpret in narrow areas such as corridors.

5.4. *Usability Study of the Proposed System.* Below we list the criterion descriptor, response mean, mode (in parentheses), and the range of responses for each criterion evaluated during the usability study:

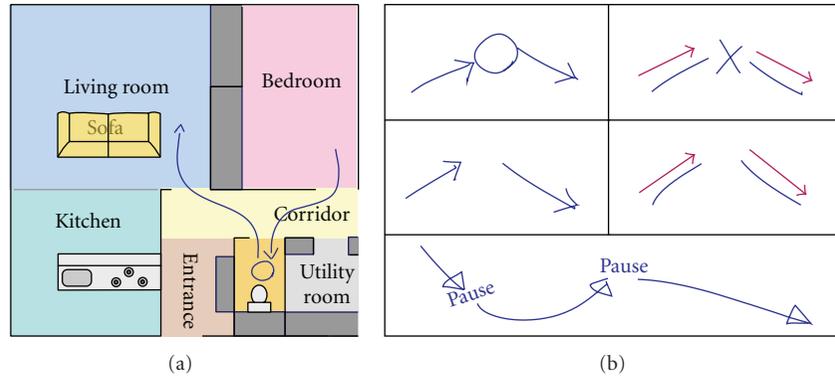


FIGURE 9: Sketch notations denoting a pause in movement.

- (i) learning to use the system: 6.375 (6) 5–7
- (ii) usefulness as a means of input: 5.625 (6) 4–7
- (iii) ease of using the system: 5.813 (6) 4–7

Most of the responses after the mode (6) were for response 7, and there was only one response at level 4 for each of the second and third criteria. The results show that the system is quite easy to learn and useful, despite the fact that it was designed without conducting a requirements study. We believe that the reason for this is the intuitive nature of querying for a locomotion pattern using a sketch.

**5.5. Feedback and Comments.** Answers to the first question of this section, “Do you think it is easier to specify a pattern of movement inside a house by using a sketch than a verbal description?” are listed as follows (the number of subjects responding in this way is indicated in parentheses).

- (i) Sketching is definitely easier (12).
- (ii) Sketching is easier in most cases (3).
- (iii) Sketching is not any easier than a verbal description (1).

While most of the subjects agreed that sketching is easier than describing verbally, the four subjects who disagreed stated that for some simple queries (such as entering a room with only one door), a textual description is easier than making a sketch.

The following are the answers to the second question “What are the other applications you would suggest for a sketch-based interface that can accept movement patterns as input?” organized in the same format as above:

- (i) searching for movement patterns outdoors (6),
- (ii) finding routes in a city (4),
- (iii) search interfaces for cities, shopping malls and so forth (4),
- (iv) support for trip planning, by including waiting times and preferred routes (3),
- (v) specifying activity in an environment (3),

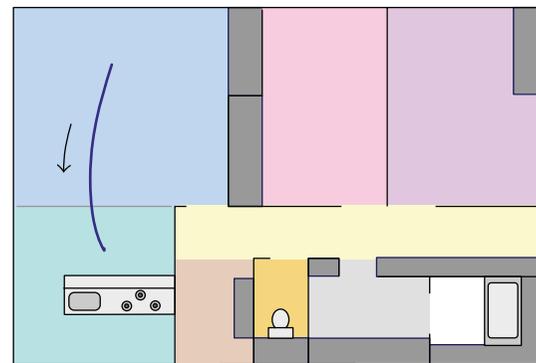


FIGURE 10: Type 2 query for movement from living room to the kitchen.

- (vi) healthcare (2),
- (vii) query for player movements in sports videos (1),
- (viii) for specifying player movements in a computer game (1).

The large number of responses was encouraging, and pointed to various applications and future research directions. Most of the suggestions are for outdoor applications, suggesting that spatial querying is more intuitive and appropriate in outdoor scenarios. For applications in indoor environments, the main restriction is the need for deployment of sensors for accurate tracking. Healthcare and surveillance are the most prospective applications, given the importance of being able to retrieve movement patterns in such applications. For outdoor applications, tracking data are easier to obtain with the availability of GPS data. While it is not possible to obtain video from very large areas, the movement patterns of vehicle fleets can easily be retrieved and shown on a map, using the proposed interaction strategy. Search in sports video is quite prospective as an immediate application in a medium-sized environment.

More than half of the subjects provided additional comments at the end of Section 5. The following are the comments most related to the main focus of this work.

- (i) It is desirable to be able to sketch paths consisting of multiple segments.
- (ii) Different body gestures and actions such as “standing, sitting, running, sleeping” should be added.
- (iii) Instead of sketching an entire path, it should be possible to sketch using a set of points.
- (iv) The system should be extended to query for activities inside the house by sketching.
- (v) It is possible to combine menu-based querying and sketch based querying, to provide users more choice.

Most of the subjects desired to see more functionality and control. Some others looked for more flexibility in entering queries, as indicated by the last comment.

## 6. Improvement of the System

*6.1. Discussion of Results.* The responses described in Sections 5.1 and 5.2 can be used to improve the user interaction strategy. It is evident that there are some common notations that the majority of the subjects agree upon. Such notations, if adopted, will make the system easier to use. According to the responses, a closed curve is the best representation for a region. Similarly, an arrow is the best candidate for representing a path of locomotion.

While most of the subjects drew arrowheads to indicate direction, the diversity of shapes and stroke-orders indicates that it will be difficult to interpret all of them automatically with sufficient accuracy. We suggest a semiautomated solution to this problem. The user can draw a line, to which an arrowhead can be added automatically. The direction the line is drawn can be used to determine the direction of the arrow. While this imposes a constraint on the direction of sketching, the results show that the users adhered to this rule even without imposing it.

The subjects did not find ambiguous or imprecise descriptions of a problem as long as the main action concerned with querying is described (Section 5.1). Most of the users were not concerned about specifying minor details of a locomotion pattern, such as pauses. While it is possible to increase functionality by adding other notations and query types to the user strategy, we believe that simplicity should be maintained so that the queries are easy to sketch.

It is evident that the notation used in the initial system for the type 1 query is not intuitive, and was actually used by only one of the subjects. This should be replaced using a closed curve, which was the choice of most of the subjects. Further, this change will make the query usable in querying for locomotion in outdoor environments that do not possess the strong and fixed partitioning present in a house.

However, the notation for Type 2 queries creates a different situation. Although none of the subjects used it, including the query type in the user interaction strategy and making them aware of it will facilitate faster querying. As demonstrated in Section 3, a sketching notation is easy to learn, and this particular notation is quite simple. Further, having this notation will not be a problem for those who do

not want to use it, as long as type 3 queries are facilitated. Therefore we consider this as a situation where a design decision can be made to support the users.

According to these observations and the results of the quantitative evaluation described in Section 3.4, the user interaction strategy and search algorithms were slightly modified to make the system more usable, and the results more accurate. The following subsections describe these changes in detail.

*6.2. Revised User Interaction Strategy.* Changes were made to the notation for the three query types, to make them more intuitive and less ambiguous. The notation for a type 1 query was changed to a closed curve, instead of scribbling used in the initial system. Using a closed curve is intuitive according to the results of the user study, and reduces ambiguities. If the area enclosed by the curve is more than that of the largest ellipse contained in the same bounding box as the sketch, the bounding box is selected as the search region. Further, if each dimension of the bounding box is within a 10% deviation from that of a room/region that it is contained in, the entire room/region is considered for searching. All the closed curves drawn by the subjects to specify a room/region confirmed to these thresholds. This strategy facilitates querying within an entire room without making a precise sketch enclosing its area, while maintaining the ability to query for arbitrary-shaped regions.

Figure 11 shows a few example queries possible under the new notation. Figure 11(a) shows a query for walking or standing in the living room. Figures 11(b) and 11(c) show how to query within a part of a predefined region, or a combination of regions and/or their parts.

For both type 2 and type 3 queries, the interface was modified to add an arrowhead automatically to strokes that are not closed curves. A query was identified as type 2 only if the sketch crosses a partition. This removes the ambiguity described in Section 5.3 (see Figure 10). This change does not impose a limitation on searching, since the modified type 1 query allows searching within arbitrarily defined regions. Type 3 queries were enhanced to facilitate querying using multiple line segments. This modification makes sketching much easier, and does not conflict with the search algorithm.

*6.3. Revised Search Algorithms.* In order to increase the precision of retrieval, the matching algorithm was modified slightly. The retrieved results were filtered to remove false positive results, using the distances between the starting points and ending points of the sketch and the retrieved path.

The normalized distance between starting points,  $D_S$  and the normalized distance between end points  $D_E$  are calculated as

$$\begin{aligned}
 D_S &= \frac{\text{Distance between } P_1 \text{ and } C_1}{\text{Length of } P}, \\
 D_E &= \frac{\text{Distance between } P_m \text{ and } C_n}{\text{Length of } P},
 \end{aligned} \tag{2}$$

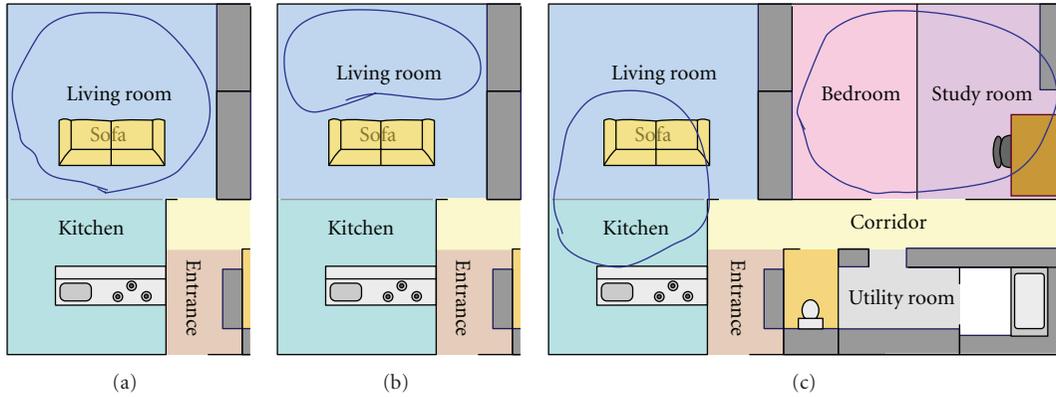


FIGURE 11: Type 1 queries using the new notation.

where  $P$ ,  $C$ ,  $m$ ,  $n$  are defined in the same notation as in Section 3.2.

A candidate path  $C$  is considered as a match only if  $D_S + D_E < 0.4$ . The value was selected empirically, to allow an average margin of error of 20% at either end, when sketching a path.

**6.4. Quantitative Evaluation.** The experiment described in Section 3.4 was repeated after making the changes, in order to evaluate the effect on performance. The precision increased from 92.5% to 95.9%, while the recall remained unchanged. The balanced F-measure increased from 95.2% to 97.3%. The ordering of the results was also improved due to the presence of a better distance measure.

## 7. Conclusion and Future Directions

We have proposed a user interaction strategy and a search algorithm for querying and retrieving from a collection of video sequences based on spatial information. These can be applied to the results of human tracking based on any type of sensors, and therefore are not restricted to environments with floor sensors. The accuracy of retrieval, when applied to real-life data from a home-like environment, was approximately 95%.

A user study was conducted to both identify how people sketch human locomotion patterns, and evaluate the proposed system in terms of usability. Changes were made where appropriate, according to the user feedback and the results of the quantitative evaluations. After the modifications, the accuracy of the system increased to about 97%, and was more intuitive to use. The user interaction strategy and search algorithms can be employed to query any type of tracking data from different environments, with minor modifications such as changes of area layouts.

Creating a formal model for the queries including time and speed will increase the versatility of spatial queries, and we are working in this direction at the moment. The feedback from the test subjects indicates that there are several prospective future directions. These include searching for locomotion patterns in outdoor environments,

querying for player behavior in sports videos, and travel planning. Work on a system for retrieval of locomotion patterns using continuously archived GPS data is currently in progress.

## Acknowledgments

The authors thank the voluntary subjects who participated in the user study, for their contribution. This work was partially supported by NICT, CREST and JST of Japan.

## References

- [1] R. Harper, *Inside the Smart Home: Ideas, Possibilities, and Methods*, Springer, Heidelberg, UK, 2003.
- [2] T. Mori, H. Noguchi, A. Takada, and T. Sato, "Sensing room: distributed sensor environment for measurement of human daily behavior," in *Proceedings of the 1st International Workshop on Networked Sensing Systems (INSS '04)*, pp. 40–43.
- [3] M. Takahashi, S. Ito, Y. Sumi, M. Tsuchikawa, K. Kogure, K. Mase, and T. Nishida, "A layered interpretation of human interactions captured by ubiquitous sensors," in *Proceedings of the 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE '04)*, pp. 32–38, 2004.
- [4] B. Gottfried, "Spatial health systems," in *Proceedings of the Pervasive Health Conference and Workshops*, pp. 1–7, November–December 2006.
- [5] Y. A. Ivanov and C. R. Wren, "Toward spatial queries for spatial surveillance tasks," in *Proceedings of the Pervasive Technology Applied Real-World Experiences with RFID and Sensor Networks (PTA '06)*, 2006.
- [6] T. Yamazaki, "Ubiquitous home: real-life testbed for home context-aware service," in *Proceedings of the 1st International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, Tridentcom*, pp. 54–59, 2005.
- [7] G. C. de Silva, T. Yamazaki, and K. Aizawa, "An interactive multimedia diary for the home," *IEEE Computer*, vol. 40, no. 5, pp. 52–59, 2007.
- [8] J. C. Augusto and C. D. Nugent, *Designing Smart Homes — the Role of Artificial Intelligence*, Lecture Notes in Computer Science, Springer, 2006.
- [9] G. D. Abowd, I. Bobick, I. Essa, E. Mynatt, and W. Rogers, "The aware home: developing technologies for successful

- aging,” in *Proceedings of the American Association for Artificial Intelligence (AAAI '02)*, 2002.
- [10] J. C. Augusto, C. D. Nugent, S. Martin, and C. Olphert, “Towards personalization of services and an integrated service model for smart homes applied to elderly,” in *Proceedings of International Conference on Smart Homes and health Telematic*, pp. 151–158.
- [11] G. Dewsbury, K. Clarke, M. Rouncefield, I. Sommerville, B. Taylor, and M. Edge, “Designing acceptable ‘smart’ home technology to support people in the home,” *Technology and Disability*, vol. 15, no. 3, pp. 191–199, 2004.
- [12] Department of Sensory Media, “Ubiquitous Sensor Room,” ATR Media Information Science Laboratories, Kyoto, Japan, <http://www.atr.jp/mis/past/im.html#b>.
- [13] R. Aipperspach, E. Cohen, and J. Canny, “Modeling human behavior from simple sensors in the home,” in *Proceedings of the 4th IEEE International Conference on Pervasive Computing*, vol. 3968 of *Lecture Notes in Computer Science*, pp. 337–348, Dublin, Ireland, 2006.
- [14] Y. Ivanov, A. Sorokin, C. Wren, and I. Kaur, “Tracking people in mixed modality systems,” in *Visual Communications and Image Processing (VCIP '07)*, vol. 6508 of *Proceedings of the SPIE*, January 2007.
- [15] A. Jaimes, K. Omura, T. Nagamine, and K. Hirata, “Memory cues for meeting video retrieval,” in *Proceedings of the 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE '04)*, pp. 74–85, 2004.
- [16] M. J. Egenhofer, “Spatial query by sketch,” *Journal of Visual Languages and Computing*, vol. 8, no. 4, pp. 403–424, 1997.
- [17] G. C. de Silva, T. Yamasaki, and K. Aizawa, “Evaluation of video summarization for a large number of cameras in ubiquitous home,” in *Proceedings of the Association for Computing Machinery Multimedia*, pp. 820–828, 2005.
- [18] G. C. de Silva, T. Yamasaki, and K. Aizawa, “Interactive experience retrieval for a ubiquitous home,” in *Proceedings of the ACM International Multimedia Conference and Exhibition*, pp. 45–48, 2006.
- [19] J. P. Chin, V. A. Diehl, and K. L. Norman, “Development of an instrument measuring user satisfaction of the human-computer interface,” in *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '88)*, pp. 213–218, 1988.
- [20] A. Blaser, “Geo-spatial sketches (A study of people’s sketching behavior and strategies),” Technical Report, National Center of Geographic Information and Analysis, University of Maine, Orono, Me, USA, 1998.



# Hindawi

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

