Situative Space Tracking within Smart Environments

Dipak Surie, Florian Jäckel, Lars-Erik Janlert
Department of Computing Science
Umeå University
Umeå, Sweden
{dipak, ens07fjl, lej}@cs.umu.se

Thomas Pederson
Innovative Communication Group
IT University of Copenhagen
Copenhagen, Denmark
tped@itu.dk

Abstract — **This paper describes our efforts in modeling and tracking a human agent's situation based on their possibilities to perceive and act upon objects (both physical and virtual) within smart environments. A Situative Space Model is proposed. WLAN signal-strength-based situative space tracking system that positions objects within individual situative spaces (without tracking their absolute positions) distributed across multiple modalities like vision, audio, and touch is presented. As a proof-of-concept, a preliminary evaluation of the tracking system was performed by two subjects within a living-laboratory smart home environment where a global tracking precision of 83.4% and a recall of 88.6% were obtained.**

I. INTRODUCTION

With the advancement in ubiquitous computing and virtual reality technologies, everyday environments like homes, airports, shopping malls, etc. are on the verge of being transformed from purely physical environments to physical-virtual environments [5] that are sensitive, adaptive and responsive to its occupants and to fulfill their immediate needs. Such environments are referred to as smart environments and are expected to be aware of its occupants and their situations. Modeling human situations is difficult considering the variety, complexity, individual differences and the dynamic nature of human situations. According to Gibson [4], human agents survive in an environment based on the information perceived by them and used for directing actions. This work briefly presents a Situative Space Model (SSM) [12] for describing a human agent's situation based on their abilities to perceive and act upon objects (both physical and virtual) within smart environments.

Tracking objects and positioning them within the situative spaces is a challenging task considering the inherent differences between physical and virtual objects [16], difficulties in establishing the boundaries of the different levels of human perception and action, and the unavailability of a suitable object tracking infrastructure that spans across vision, audio and touch modalities. Detecting human situations and behavior using video and audio tracking in a smart home environment is described in [13]. Such an approach requires multiple cameras and microphones, which are often considered intrusive on the occupant's privacy. Grocery object recognition using a wearable camera is described in [20]. The privacy issue is addressed but the recognition is restricted to ambient or static objects while visual tracking issues like motion and blur, illumination variation, etc. still remain to be solved. Electrooculography based eye gaze tracking [17]; object manipulation tracking based on passive RFID [21] and tiny wireless sensors [8]; accelerometer based gesture tracking [22]; etc. are interesting related works, but they focus on specific perception and/or action information channels, while we are interested in a general infrastructure for tracking the situative spaces. Localizing physical objects using received signal strength indicators (RSSIs) and vibration data from active RFID tags, complemented by sensors embedded in the environment to detect human behaviors is described in [15]. Since smart environments usually have wireless networks, signal-strength-based approaches for localization [1] could be considered advantageous (no need for additional hardware). We present an approach for tracking and positioning objects within the situative spaces using WLAN signal-strength measures complemented by the objects' characteristics.

Situative space tracking is useful for facilitating situated interaction where information about the usage situation is integrated within human-computer interaction (HCI); for understanding and exploring the natural mapping between a human agent's perception and actions; for recognizing human activities as shown in our previous work [9]; for executing context-aware services [10]; and for integrating the physical and the virtual in mixed-reality environments [11, 16, and 5]. This paper is arranged as follows: section 2 describes a *Situative Space Model* and the operational definitions of its components; section 3 presents a WLAN signal-strength-based *situative space tracking system* including its architecture; section 4 presents the experimental setup, evaluation results and a conclusion.

II. A SITUATIVE SPACE MODEL (SSM)

The perspective in HCI research on smart environments is moving towards conceiving human beings as mobile agents in a dynamically changing environment populated by physical and virtual objects alike considering their situatedness for facilitating better HCI experience within such environments. Modeling human behavior within a smart environment using a situation model consisting of the environment, users, and activities is described in [14]. Our work (the situative space model [12]) focuses on capturing what a specific human agent can perceive and not perceive, affect and not affect at a given moment in time. According to psychological studies [18], there is a relationship between object perception and recognition, action understanding, and action execution. While perceiving and recognizing objects is important for action execution, the converse might also be true [19]. The SSM is also inspired by the proximity principle which states that "*things that are close tend to matter; things that matter tend to*

978-0-7695-4149-5/10 $26.00 © 2010 IEEE
DOI 10.1109/IE.2010.35

be(come) close". Objects (including other agents) that are close to a human agent have a fair chance of getting their attention and figure in the agent's current cognitive processes and activities compared to objects at a distance.

A. An Egocentric Perspective

Human beings have a self-centered perspective on their environment before any other perspective, such as an other-person perspective. The world is so large and so rich in details that any agent with limited cognitive capacity must necessarily narrow its focus in some manner. Human beings have physical bodies that are located at a single particular place and oriented together with their limbs and sense organs in particular directions at any particular time. That gives them a natural primary vantage point for selecting which details and aspects to attend to (perception) and also gives them a focus for actions given their current bodily situation in the world.

The term "egocentric" has been chosen to signal that it is the human body and mind of a specific human individual that serve as center of reference to which situation modeling is anchored. The term "egocentric" should not be taken as a synonym for "selfish" but instead as a lower-level approach which human agents in general are forced to adopt to perceive and act in the world based on their senses and cognitive abilities, even when working in groups and with shared goals.

B. Components of a Situative Space Model

Within smart environments, the classical HCI concepts of input and output needs to be substituted with something that works for both physical and virtual object manipulation. Also, since action is often inseparable from perception in performing physical activities, we propose an agent-centered SSM that varies continuously with the human agent's movement of body and body parts, and is based on their perception and action possibilities. Operational definition of the components of this model is as follows (also refer to Fig. 1):

- **World Space (WS):** A space containing the set of all physical and virtual objects to be part of a specific situative space model.
- **Perception Space (PS):** The part of the space around the agent that can be perceived at each moment. Perception space can be given a simple geometrical interpretation like a cone in the case of vision (refer to Fig. 1) as a rough approximation. Objects may occlude other objects and thus create (temporary) holes in the space.
- **Recognizable Set (RS):** The set of objects currently within perception space that are within their recognition distances, i.e. it is possible for a human agent to recognize what type of objects they are.
- **Examinable Set (ES):** The set of objects currently within perception space that are within their examination distances, i.e. it is possible for the human agent to not only recognize the objects but also their states. Normally, we expect the Examinable Set to be a proper subset of the recognizable set.
- **Action Space (AS):** The part of the space around the agent that is currently accessible through the agent's actions. Objects within this space can be directly acted on. The outer range limit is basically determined by the physical or virtual reach of the agent, but obviously depends qualitatively also on the type of action and the properties of objects involved; e.g., a reachable object may be too heavy to handle with outstretched arms. Since many actions require perception to be efficient or even effective at all, action space is qualitatively affected also by the current shape of perception space.
- **Selected Set (SS):** The set of objects currently being physically or virtually handled (touched, gripped; or selected in the virtual sense) by the agent.
- **Manipulated Set (MS):** The set of objects whose states (external as well as internal) are currently in the process of being changed by the agent. Normally, we expect the manipulated set to be a subset of the selected set.

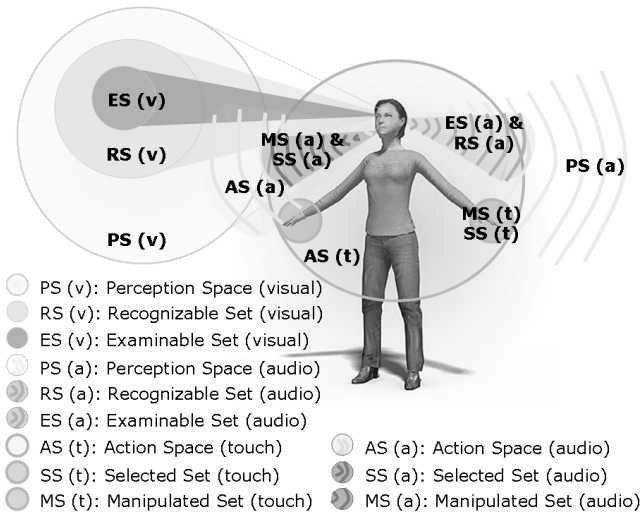

Figure 1. A Situative Space Model spanning across multiple modalities.

Human agents perceive and act not based on a single modality but through multiple modalities that interact with each other to create a unified representation of perception and action. This introduces a need to let the situative space model span across multiple modalities, but also to operationalize mechanisms for multimodal fusion. The situative space tracking system proposed in this paper will focus on tracking the perception space (visual and audio modality), recognizable set (visual modality), and action space (touch modality). The sensing of selected and manipulated sets using simple state change sensors is described in our previous work [8]. Note that in Fig. 1, PS (v) refers to the perception space along the visual domain while PS (a) refers to the perception space along the audio domain.

C. Establishing the Situative Space Boundaries

To operationalize the situative space model, mechanisms to establish the situative space boundaries are needed. The proposed model is a simplification with potential improvement possibilities. In addition to the type of errors that may occur due to imperfect modeling of the human agent's perception and action possibilities, errors of the situative space model itself,

there will be errors that occur due to the imperfections of the technological infrastructure used for tracking the situative spaces. At this time the accuracy of available, feasible technologies is inadequate, but development in this area is rapid.

III. SITUATIVE SPACE TRACKING SYSTEM

Traditionally, physical environments contain physical objects, but smart environments contain objects that could provide virtual functionalities without compromising their physical utility. Such augmented physical objects are referred to as "smart everyday objects" (SEOs). In Fig. 2, a bathroom mirror is used by a human agent to view his physical appearance and in parallel check his day schedule (which is a virtual object) through a LCD display (which is a mediator [5]) equipped in the bathroom mirror SEO. Mediator is a common name given to objects (like sensors, actuators, input/output devices, interaction/interface devices, user interfaces, etc.) that mediate events between the physical and the virtual world. Mediators are expected to be both perceptually and cognitively transparent. This allows physical and virtual objects to be co-present in time and space from a human agent's perspective. Some plain physical objects (for instance, a wash basin below the bathroom mirror SEO) utilize their spatial co-location with SEOs to become an inherent part of the smart environment. This approach eliminates the need to augment all physical objects with technology and there by indirectly addresses the issue of scalability of SEOs in a smart environment.

The situative space tracking system is distributed across a set of SEOs and a wearable personal server [6]. The SEOs are equipped with ASUS WL-167g WLAN adapters and thin-client boards for wirelessly communicating with an off-the-shelf wireless access point (WRT54GL) and a customized directional antenna embedded to the personal server. The antenna is shielded with aluminum to make it directional. The personal server is expected to be a part of a human agent's wearable outfit, but a notebook PC is used for the initial prototype. Java Wireless Research API from Luleå University of Technology is used by the individual SEOs to calculate the WLAN signal strength measures that indicate their proximity with reference to a human agent's body [1]. Since the SSM itself is developed based on a human agent's proximity to surrounding objects, a proximity tracking infrastructure seems appropriate. The off-the-shelf radios embedded in the SEOs serve both as an object-tracking infrastructure and for data communication. Since indoor walls play an important role in determining the boundaries of the situative spaces, all indoor walls were embedded with a layer of aluminum foil to dampen the signal strength measures of SEOs behind the walls. The situative space tracking system does not provide absolute proximity measures between a human agent and a set of SEOs, but instead finds the most probable situative space for the individual SEOs. To calculate absolute proximity measures, additional information about

Figure 2. Components of a Cutting Board Smart Everyday Object (SEO).

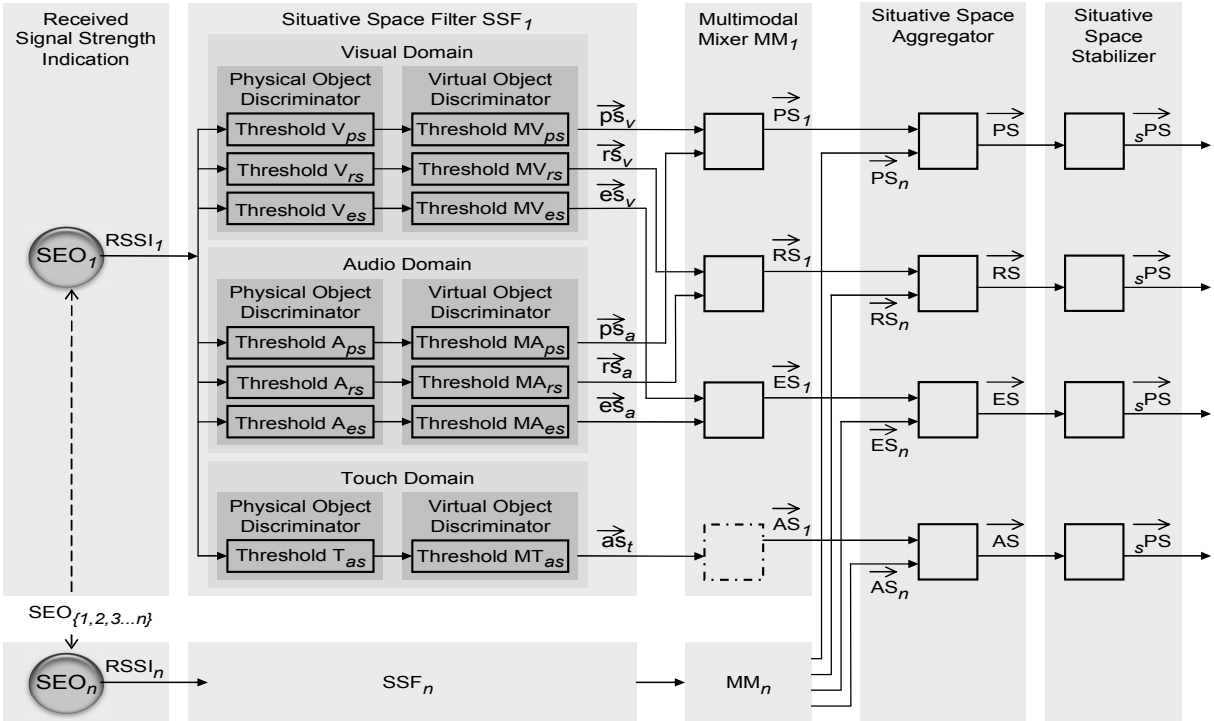

Figure 3. An architecture for situative space tracking within smart environments.

indoor walls, signal propagation characteristics, etc. are required, while our application does not require absolute proximity measures. Fig. 3 shows the architecture for the situative space tracking system comprising five components. A feed-forward model is used for simplicity. Received Signal Strength Indicator, Situative Space Filter and Multimodal Mixer are components implemented in the individual SEOs. Situative Space Aggregator and Situative Space Stabilizer are part of the personal server. Low-level calculations by the SEOs reduce the wireless bandwidth requirements in exchanging data with the personal server within a star network topology. SEOs already possess computational power to, for instance display virtual visual and audio objects, it justifies the system design where low-level calculations are handled by the SEOs.

A. *Received Signal Strength Indicator*

Detecting proximity based on received signal-strength indicator (RSSI) is highly sensitive to scattering, reflection, signal attenuation and measurement noise [2]. Moreover, RSSI values fluctuate over time and the stability of signal-strength measures varies among different SEOs. Signal patterns of the different SEOs were collected during an initial pilot study (refer to section IV.A). This study was useful in establishing simple statistical techniques to discard measures that are inappropriate and inconsistent over time. Information about factors that influence signal propagation within smart environments like indoor walls, influence of other objects, etc. are ignored since the system is intended for tracking the objects' semantic location and not their physical location. $SEO_{\{1,\ 2,\ 3\ \ldots n\}}$ represents the set of SEOs within a smart environment. The SEOs measure signal strengths at 5 Hz, and apply appropriate statistical methods to remove inconsistent signals before sending the average RSSI measure to the situative space filter every 2 seconds.

B. *Situative Space Filter*

SEO_1 for instance, might be perceived by a human agent through multiple modalities, introducing a need to filter the physical object (and co-located physical objects), virtual object(s) and mediator(s) making up the SEO_1 across multiple modalities. Depending upon the modality, other characteristics of the object affect its position within the situative spaces. One such characteristic is the object size along the visual modality. Bigger objects could be perceived from a distance compared to smaller objects. Physical objects do not vary their size (if we ignore exceptional cases) with ease, in contrast to virtual objects. Virtual object characteristics like object loudness or object size could be varied dynamically and the cost involved in such changes are often negligible. This makes it difficult for the situative space tracker to position the virtual objects within the situative spaces. At present, we have experimented with virtual objects that do not change their characteristics dynamically, however future versions should handle such dynamic changes.

Refer to Fig. 3. The situative space filter SSF_1 considers the $RSSI_1$ value across the visual, audio and touch domains separately. Within the individual domains, the $RSSI_1$ value is first fed through a Physical Object Discriminator to position the physical object (and co-located physical objects) within the situative spaces. Thresholds like V_{PS} (Visual Perception Space), V_{RS} (Visual Recognizable Set), V_{ES} (Visual Examinable Set), A_{PS} (Audio Perception Space), A_{RS} (Audio Recognizable Set), A_{ES} (Audio Examinable Set), and T_{AS} (Touch Action Space) are used. The $RSSI_1$ value is then fed through a Virtual Object Discriminator to position the virtual object(s) within the situative spaces. This is done in two steps where the mediator(s) within the SEO_1 is first positioned within the situative spaces based on thresholds like MV_{PS} (Mediated Visual Perception Space), MV_{RS} (Mediated Visual Recognizable Set), MV_{ES} (Mediated Visual Examinable Set), MA_{PS} (Mediated Audio Perception Space), MA_{RS} (Mediated Audio Recognizable Set), MA_{ES} (Mediated Audio Examinable Set) and MT_{AS} (Mediated Touch Action Space). Then the virtual objects characteristics like object size (for visual virtual objects), object loudness (for audio virtual objects), etc. are mapped to the mediators' positions to finally position the virtual objects within the situative spaces. In the future, additional contextual information like background light, background noise, etc. could be used to further filter objects within the situative spaces. The situative space filter SSF_1 outputs the following vectors ps_V, rs_V, es_V, ps_a, rs_a, es_a, and as_t that contain information about the position of the physical objects, virtual objects and mediators contained by SEO_1.

C. *Multimodal Mixer*

The Multimodal Mixer fuses the objects spread across different modalities within individual situative spaces. Such a fusion is motivated by a human agent's ability to perceive or act upon an object through multiple modalities. Physical objects to some extent are easier to handle computationally compared to virtual objects that might possess multiple manifestations. It is important for the multimodal mixer to effectively complement the natural characteristics of individual objects (i.e. some objects are characterized by their manifestation modality like an alarm clock with audio modality or a wall clock with visual modality) and at the same time use redundancies as a way of improving the robustness (i.e. eliminate noise and sensing errors). Issues of correlation between modalities that enhances or affects the perception and action possibilities of individual objects should be considered by the multimodal mixer. For instance, an object that is both visually and aurally perceivable could be given higher probability of being in the fused perception space compared to an object that is only aurally perceivable. Welch & Warren [3] propose the hypothesis that processes involving multiple modalities follow modality appropriateness, where visual modality dominates over audio modality or touch modality for spatial tasks, while audio modality dominates for temporal tasks. Multimodal Mixer MM_1 for instance fuses Vector ps_V and ps_a to obtain Vector PS_1 that represents the set of physical objects, virtual objects, and mediators within the perception space.

D. *Situative Space Aggregator*

The Situative Space Aggregator is responsible for gathering situative space information from different SEOs that are in the close proximity to a human agent. Vectors PS_1 to PS_n for

instance are aggregated to Vector PS representing the global Perception Space. During the aggregation process, relationship among objects (both physical and virtual) could be used to improve accuracy and to compensate for the lack of mediators in some physical objects. Co-location of objects could be considered along both spatial and temporal domains. Magic Touch [5] is an object location tracking system based on RFID and ultrasound technologies where the concept of containment is used to co-locate objects within a container with the container itself. For instance, if papers are within a folder and the folder is tracked in a new location, then the papers are also assumed to be in the new location. The co-location relationship among objects was hard-coded during the initial pilot study. This approach suffers from obsolete relationship problems in a dynamic environment where the objects are not stationary, but removes the need for embedding technology within simple physical objects, especially if they are stationary.

E. Situative Space Stabilizer

Human agents are often mobile or perform body movements that change their orientation towards objects within smart environments. This creates scenarios where objects enter and leave the situative spaces dynamically and sometimes even too fast for a human agent to perceive or act upon them. Human agents do visual stabilization of the objects that enter and leave their visual perception space. Their visual system makes use of the fact that physical objects do not flicker in and out within physical environments. Hence, the role of the situative space stabilizer is to stabilize the objects within the situative spaces by checking their presence or absence over a period of time. Our hypothesis is that such an approach ensures that an object for instance, within the perception space is actually perceivable by a human agent. The stabilizer also removes noise due to signal fluctuations that are common in such environments, thereby making sure that only stable objects are positioned within the individual situative spaces. The situative space stabilizer introduces time domain to the situative space model. The stabilization time is to some extent dependent on the application, and also the objects involved. For instance, a virtual video object might impose larger stabilization time compared to a simple image object. Vector $_S$PS for instance, represents information content within the perception space that is stabilized over a period *S*.

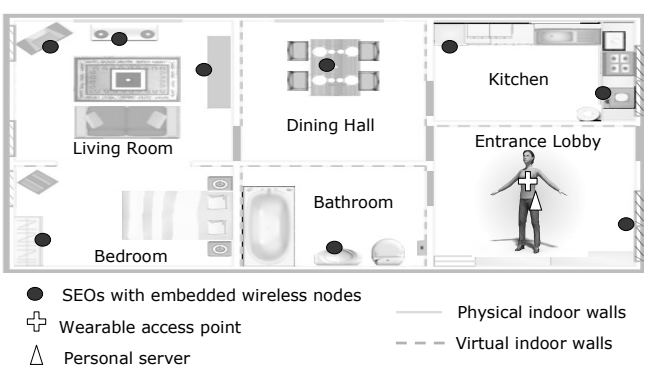

Figure 4. A living-laboratory smart home environment populated with SEOs.

IV. EVALUATIONS AND CONCLUSION

A living-laboratory smart home environment [7] populated with 9 SEOs representing a set of physical objects (and co-located physical objects), mediators and virtual objects was used for evaluating the situative space tracking system. The distance between the 2 farthest SEOs was 8.5m (approx.) and the 2 closest SEOs was 0.9m (approx.). The smart home environment is intended for single inhabitants. Refer to Fig. 4. Multi-agent environments pose additional challenges like situative space sharing, resolving conflicts, etc. which are not addressed in this paper. The two subjects that took part in the experiments provided the ground truth by calling out loud, while hand-written notes were made by an observer. To understand the system decisions, time-stamped log files were obtained from the individual SEOs and the personal server. The subjects had some experience of using computers. They were given a brief introduction to our work, and the concepts of SEOs, perception, action, etc. were described. Windows Media Player 9.0 was used as the virtual environment containing a variety of virtual objects like text objects, audio objects, video objects, image objects, etc. that are distributed across multiple modalities. The experimentation took place in two phases. During the first phase, the situative space thresholds were established, while during the second phase the situative space tracking accuracy was established.

A. Coverage of the Situative Space Tracking System

The Transmission-Reception (Tx-Rx) range is important to make sure that the proposed tracking system has 100% coverage within the smart home environment. During an initial pilot study, WRT54GL access point's default transmission power of 42 mW was established appropriate considering a Tx-Rx range of 72 meters (with -65 dB signal strength as a threshold to maintain network quality) in the smart home backyard (outdoor environment). However, such an evaluation is of lesser importance compared to their performance indoor. A Tx-Rx range of 48.27 meters was obtained as an average (among 10 readings) with a single indoor wall obstruction within the 54 m^2 smart home environment, while the network still showed acceptable quality for multiple indoor wall obstructions.

B. Signal-Strength-based Thresholds for the Situative Spaces

The subjects positioned themselves in 32 different locations in the smart home environment. They spent 6-8 minutes at each location, constantly changing their proximity and orientation with reference to the surrounding SEOs. Every 5-10 seconds, the subjects told the experimenters in what set or space specific physical objects or mediators were present, according to their own perception and action possibilities. This information was used to define the signal strength thresholds (refer to Fig. 5) at the border of the spaces and sets by identifying the time-stamped values recorded in log files generated by the WLAN adapters in the corresponding situations. Even though this study was a preliminary one, it was useful to calculate the situative space thresholds as a first step in the exploration process. The variation between the subjects was not considered in calculating personalized thresholds (simple threshold average of the two subjects was

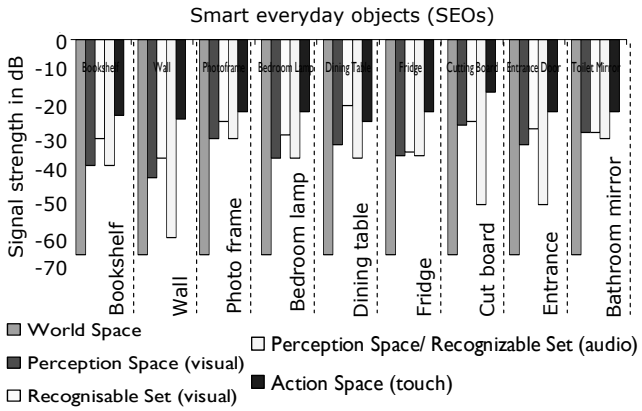

Figure 5. Signal-strength-based situative space thresholds for individual SEOs.

used), but would be useful to consider in personalizing the situative space model (and indirectly the smart environment) to individuals in the future.

C. *Precision and Recall Values*

The situative space tracking system was tuned to the situative space thresholds in Fig. 5. The same subjects performed everyday activities in the smart home environment, reporting objects perceived as being inside the spaces and sets every 10-20 seconds. Using that information as ground truth, and by comparing it to the log files generated, the system's accuracy was obtained as shown in Table 1. The subjects preferred to provide the ground truth while being stationary compared to while being mobile. Since the intensity of interaction between a human agent and their environment is low while being mobile, the influence of the absence of proper ground truth in such situations could be considered negligible. The everyday activities resulted in constant changes within the situative spaces [9]. Even though the study is limited in the number of subjects and the duration spent by each subject which was about 60 minutes each, the results encourage further research in this direction. The system performs well in determining perception space (precision of 99.80%) and recognizable set (precision of 99.78%) in the visual domain. System performance in the audio domain is worse (precision of 63.62%) due to background noise and the fact that subjects were not confident enough in distinguishing audio information within the situative spaces. For the action space, the precision reached 100% but a recall value of 55.03% meaning that the system was unable to make a decision 44.97% of the time, which is unacceptably high and needs to be reduced in the future.

Table 1. Precision and recall values of the situative space tracking system.

	True Positive	False Positive	False Negative	Precision (in %)	Recall (in %)
Perception Space (visual)	743	3	16	99.6	97.9
Perception Space (audio)	717	410	86	63.6	89.3
Recognizable Set (visual)	450	1	31	99.8	93.6
Action Space (touch)	164	0	134	100.0	55.0
Global	2074	414	267	83.4	88.6

Thus our attempt to validate a theoretically-grounded SSM in practice using WLAN signal-strength-based tracking is presented.

REFERENCES

[1] P. Bahl and V. Padmanabhan. RADAR: An in-building RFbased user location and tracking system. *IEEE Infocom*, 2000.

[2] H. Hashemi. The indoor radiation propagation channel. *Proceedings of the IEEE*, 81(7):943–968, 1993.

[3] R.B. Welch, D.H. Warren. Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.* 88, 638–667, 1980.

[4] J. J. Gibson, *The Ecological Approach to Visual Perception*, 1st ed. Lawrence Erlbaum Associates, October 1986.

[5] T. Pederson. From Conceptual Links to Causal Relations — Physical-Virtual Artefacts in Mixed-Reality Space, PhD thesis, Dept. of Computing Science, Umeå university, report UMINF-03.14, 2003.

[6] R. Want, T. Pering, G. Danneels, M. Kumar, M. Sundar, J. Light. The Personal Server: Changing the Way We Think about Ubiquitous Computing, Ubicomp'02, 2498, January 2002.

[7] S. S. Intille, K. Larson, J. S. Beaudin, J. Nawyn, E. M. Tapia, P. Kaushik. A living laboratory for the design and evaluation of ubiquitous computing technologies, *CHI '05 extended abstracts on Human factors in computing systems*, ACM, 1941-1944., 2005.

[8] D. Surie, O. Laguionie, T. Pederson. Wireless Sensor Networking of Everyday Objects in a Smart Home Environment, 4th International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Sydney, Australia, 2008.

[9] D. Surie, T. Pederson, F. Lagriffoul, L-E. Janlert, D. Sjölie. Activity Recognition using an Egocentric Perspective of Everyday Objects, Proc. of International Conference on Ubiquitous Intelligence and Computing, Hong Kong, China, Springer LNCS 4611, 246-257, 2007.

[10] A. Dey. Understanding and Using Context, Personal and Ubiquitous Computing, Vol. 5, No. 1, 4-7, 2001.

[11] P. Milgram, A. F. Kishino. Taxonomy of Mixed Reality Visual Displays, IEICE Trans. on Information and Systems, E77-D(12), 1321-1329, 1994.

[12] T. Pederson, L-E. Janlert, D. Surie. Setting the Stage for Mobile Mixed-Reality Computing - A Situative Space Model based on Human Perception, IEEE Pervasive Computing, 2010.

[13] O. Brdiczka, M. Langet, J. Maisonnasse, J.L. Crowley. Detecting human behavior models from multimodal observation in a smart home, IEEE Trans. on Automation Science and Engineering, 6 (4), 588-597, 2009.

[14] O. Brdiczka, J.L. Crowley, P. Reignier. Learning situation models in a smart home, IEEE Trans. on Systems, Man and Cybernetics, special issue on Human-Centered Computing, 39 (1), 56-63, 2009.

[15] M. Li, T. Mori, H. Noguchi, M. Shimosaka, T. Sato. Use of active RFID and environment-embedded sensors for indoor object location estimation, 3rd Int. Universal Communication Symposium, ACM, 2009.

[16] D. Surie, T. Pederson, L-E. Janlert. The easy ADL home: A physical-virtual approach to domestic living, JAISE, 2010.

[17] A. Bulling, D. Roggen, G. Tröster. Wearable EOG goggles: eye-based interaction in everyday environments, *Extended Abstracts on Human Factors in Computing Systems,* CHI '09, ACM, 3259-3264, 2009.

[18] K. Nelissen, G. Luppino, W. Vanduffel, G. Rizzolatti, G. Orban. Observing Others: Multiple Action Representation in Frontal Lobe, Science, Vol. 310, 332-336, 2005.

[19] L.L. Chao, A. Martin. Representation of Manipulable Man-Made Objects in Dorsal Stream, NeuroImage, Vol. 12, 478-484, 2000.

[20] M. Merler, C. Galleguillos, S. Belongie. Recognizing groceries in situ using in vitro training data, In *CVPR*, 1–8, 2007.

[21] M. Philipose, K. Fishkin, M. Perkowitz, D. Patterson, D. Fox, H. Kautz, D. Hahnel. Inferring Activities from Interactions with Objects, *IEEE Pervasive Computing,* 50-57, 2004.

[22] J. Kela, P. Korpipää, J. Mäntyjärvi, S. Kallio, G. Savino, L. Jozzo, D. Marca. Accelerometer-based gesture control for a design environment, *Personal and Ubiquitous Computing,* 285-299, 2006.