Regular Paper

Estimating Interruptibility in the Home for Remote Communication Based on Audio-Visual Tracking

Yoshinao Takemae,† Takehiko Ohno,† Ikuo Yoda†
and Shinji Ozawa††

This paper presents a method for automatically estimating human interruptibility in the home environment. To make online remote communication smoother, determining if it is appropriate to interrupt the remote partner is critical. As a first step in achieving this goal, several audio-visual features, extracted from data streams captured by a camera and a microphone, are correlated to human assessments of own interruptibility. Based on these features, the level of interruptibility is estimated using the trained Support Vector Regression (SVR) technique. Finally, we discuss the potential of our method based on the results of several experiments.

1. Introduction

With the growth of recent broadband networks such as optical fiber and ADSL, online remote communication services such as the videophone and teleconferencing are becoming as popular as telephones and cellular phones. The next-generation network must provide advanced online remote communication services for realizing smooth and efficient communication like face-to-face communications.

In face-to-face communications, humans find it easy to start communication, i.e., to interrupt someone, because we can determine the partner's *awareness information*, which is defined as presence, surroundings and activities, and nonverbal cues like gaze, facial expression and gestures etc.[1] However, remote communication prevents us from easily initiating communication since really effective feedback is not available until the partner accepts the call. For example, most of us have regretted calling someone at the wrong time and wished that we had known that the time was inappropriate before we called.

One means of conveying some level of awareness information about the intended partner is to set a video camera at the partner's site and leave it on continually. This approach, which directly conveys visual information captured by the camera is, however, unacceptable since it leaks private information. The need to ensure

human privacy is paramount to gain user acceptance.

Our goal is to support remote communication so that people can start communication when it is appropriate for the partner. To this end, it is essential that the interruptibility of the partner be automatically acquired and be represented without leaking private information. Using this information allows the user to determine the most appropriate timing to contact the partner intended.

In order to advance research towards this goal, this paper focuses on the home environment like the living room and proposes an automatic method that estimates the level of human interruptibility based on audio-visual tracking. The rest of this paper is organized as follows: Section 2 summarizes related work. Section 3 introduces our data on human interruptibility. Section 4 presents our method. Section 5 describes the experiments conducted to evaluate our method. Section 6 shows the challenge and future works for extending this work to create a more practical system. We summarize the paper in Section 7.

2. Related Works

This section describes related works.
2.1 Studies on Interruptions and Availability

Although interruption was first studied in the field of classical psychology, it has also been the focus of recent HCI research with the view of realizing design guidelines for deciding when to interrupt people in workspaces [2].

Hudson, et al. investigated key human be-

---

† NTT Cyber Solutions Laboratories, NTT Corporation
†† Department of Information and Computer Science, Keio University

haviors that might be related to human interruptibility in private offices (working environments) [3]. Based on their analysis of the results collected, they found that the presence/absence of speech was a significant indicator of the level of interruptibility. However, since they manually tagged voice and video streams, this method has not been implemented as a practical system.

Minakuchi, et al. proposed a method for automatically estimating the user's interruptibility at his/her desk [4]. This method assigns one of three levels of activity according to the usage of pens, the presence of sound, keystrokes, and the movements of the mouse cursor. Experiments showed that the resulting estimations were 43.8% accurate.

Horvitz, et al. proposed a method for inferring the cost of interrupting a user working at his desk, based on multiple information such as audio-visual tracking, on/off of typing, kind of application used and scheduler entries etc.[5]

Avrahami, et al. focused on Instant Messaging (IM) which has been growing in popularity for personal and work related communication [6]. IM systems result in messages often arriving at inconvenient or disruptive times. To solve this problem, they proposed a statistical model that uses features like the day of week, hour, and the kind of application being used. It can successfully predict responsiveness to incoming instant messages, that is, whether the receiver is likely to respond to a message within a certain time period.

While most conventional studies including the works mentioned above, explored the issues of interruption and availability in workplace environments [3]~[8], very little research has been done on human interruptibility in the home environment.

2.2 Tracking Human Activity in Home Environment

With the recent growth of ubiquitous computing techniques, many works such as *Aware Home* [9], *Ubiquitous Home* [10], and *EasyLiving* [12] have attempted to make human everyday life easier and richer in the home. These studies used devices such as cameras, microphones, and proximity sensors to detect human activity as robustly as possible. However, most works considered only simple actions like sitting and standing. Very little research has been achieved in detecting the more complicated attribute of human interruptibility.

3. Data Collection and Analysis

The final goal of our research is to develop a method that can automatically estimate interruptibility of the communication partner when she/he is anywhere within the home. As a first step, this paper focuses on the living room for the following reasons: 1) In the household tested, the housewife (subject) did not have her own cellular phone and the fixed phone was set in the living room. 2) Prior investigations (for two weekdays) found that the housewife was most often in the living room, some 39% of the total time.

Based on the above ideas, in order to investigate the characteristics of human interruptibility in the living room, this section describes the capture and analysis of data associated with human interruptibility. Note that this paper also focused on the case where multiple subjects shared the living room in addition to the case where only one subject was in the room, because when the multiple people were interacting such as speaking and eating, the interaction impacted their interruptibility.

3.1 Capture System

We constructed a PC-based audio-visual capture system, and placed it in the living room of a Japanese family of four. The Windows PC was connected to a USB camera, a microphone, and a ten key pad. The camera had a wide-angle lens and was attached to the ceiling. The gray scale output image (Mpeg format) had a resolution of 320×240 pixels at about 25 [frame/s], as shown in **Fig. 1**. The directional microphone was set on a wall and directed at the center of the table for capturing

Fig. 1 Examples of a captured image and ten key pad.

only the occupants' speech behavior (the TV sounds were also captured). 16-bit mono audio data was recorded at 44 [KHz]. The ten key pad, used by the subjects to input the level of interruptibility, was set on a corner of the table. Recording was performed from 9 am to midnight for about 3 weeks.

3.2   Subjects

We chose two subjects who often picked up and responded to telephone calls to the family. One was a 56-year old housewife. She sometimes went out shopping and dining with close friends. The other, a 22-year old daughter, had a weekday job so she often went out. These subjects were paid. Note that the father and son had weekday jobs and were often outside during the daytime.

3.3   Procedures

We focused on the collection of data of human interruptibility . In order to capture natural data, subjects were asked to live life as usual. A vibration timer with repeat function was attached to each subject as shown in **Fig. 2**. This timer vibrated every 15 minutes. If the timer vibrated when the subject was in the living room, the subject entered her current level of interruptibility (as if a phone call from remote human was received) as soon as possible with one level from a 5-point scale, and then selected her current behavior such as eating and drinking, speaking, sleeping, reading, writing, watching TV, listening to music, cleaning the room, folding the washing, and relaxing. **Table 1** shows the concrete meaning of each levels, which was given as instruction to the subjects. We eliminated the data collected in the morning of the first day since it was taken to cover a training period. The number of self-reported data entries from the housewife and the daughter was 348 and 49, respectively. We collected a total of 397 data points from the two subjects.

3.4   Characteristics of collected data

This section shows the results of our analysis of the human interruptibility data collected.

3.4.1   Overall distributions of Interruptibility Self-reports

Figure 3 shows the overall distributions of interruptibility for the two subjects. Based on these distributions, the least (most)-

Fig. 2   One subject with vibration timer.

Table 1   Instructions to subject for entering human interruptibility.

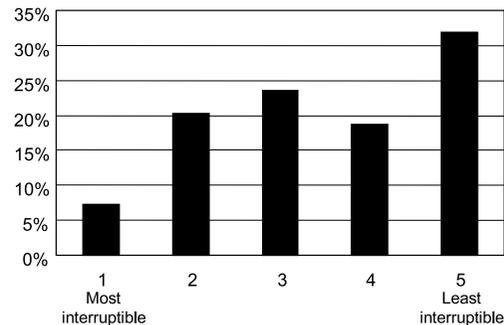| Levels	Meanings
1	You can freely receive the phone call. This is lowest irritation at being interrupted.
2	You can receive a phone call.
3	You can not tell which level your interruptibility is.
4	You hesitate to receive a phone call.
5	You do not want to receive a phone call. This is highest irritation at being interrupted.

Fig. 3   Overall distribution of interruptibility (self-reported) of two subjects. Level 1 (5) shows most (least) interruptible.

interruptible level was selected 7.3 (31.5)% of the time. As expected, there were only a few occasions when the subject wanted to receive a call.

3.4.2   Estimating Human Interruptibility from Human Behavior

We investigated how human behavior was related to human interruptibility in a living room by obtaining clues about the audio-visual features that could be used to estimate human interruptibility. The main results are as follows.

- Speaking. 91.5% of the 147 samples indicating this activity were distributed from level 3 to 5.

Another approach is for the subjects to input the level of human interruptibility while reviewing a video with voice. Unfortunately, this approach places excessive loads to the subjects. Accordingly, this paper used the online approach.

- Eating and drinking. 100% of the 20 samples indicating this activity were distributed from level 3 to 5.
- Sleeping. 100% of the 34 samples indicating this activity were assigned to only level 5.
- Relaxed. 100% of the 28 samples indicating this activity were assigned to only level 1.

4. Proposed Method

This section introduces the proposed method; based on audio-visual tracking, which can automatically estimate human interruptibility.

4.1 Preconditions

Instead of directly extracting human behavior such as sleeping, relaxed, eating and dirking, which were related to based on the observation results shown in Section 3.4.2, we focused on indirect audio-visual low-level features, considering the following requirements for designing a method that can recognize human interruptibility: 1) This work aims at supporting remote communication so that the user can start communication when it is appropriate for the partner intended, by conveying the level of interruptibility of the partner to the user beforehand. For real-time and online applications like this, computation cost must be minimized as much as possible. 2) It is difficult for current vision-based techniques to accurately recognize human behavior like relaxing, sleeping. Even if this is possible, advanced techniques such as facial expression and emotion analysis needed. This demands many high resolution cameras set in the room observed so that occupants' facial images are always tracked and captured. For this reason, system cost would be prohibitive. Computation costs would also be high. 3) We aim at implementing the proposed method as simply as possible. To this end, it is desirable to restrain device cost. This paper used only a cheap USB camera with wide-angle lens and microphone. The camera is set to the ceiling in order to capture the occupants' behavior throughout the living room.

4.2 Overview

Figure 4 overviews our method. It consists of three online modules: data gathering, feature extraction, and interruptibility estimation. In data gathering, a camera captures visual data of the occupants in the room (See Fig. 1). A microphone captures the occupants' audio data. Feature extraction extracts several audio-visual features such as voice power and occupants' mo-
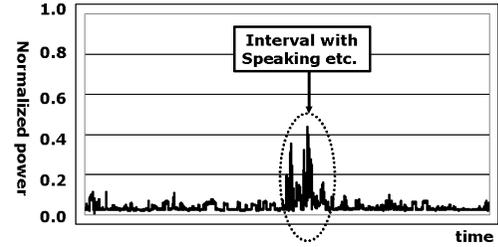

Fig. 4 Overview of proposed method.

Fig. 5 Example of powers of captured voice signals. Note that the powers have been normalized.

tion. Interruptibility estimation uses a trained Support Vector Regression (SVR) system [14] to output the level of human interruptibility for the room targeted.

4.3 Feature Extraction

To create a really practical system for online remote communication, we focused on the audio-visual low-level features that could be extracted with relatively little delay as follows.

4.3.1 Audio Features

Based on findings of past research [3] and our analysis, we found that talking with some one else in the room is an important indicator of irritation with an interruption. We selected the following audio features to extract this behavior.

(A1) Voice Power

First, the logarithmic power of the recorded audio signal in 0.1 [second] intervals over the past N [seconds] at each time of self-report entry is calculated, as shown in **Fig. 5**. Next, the average and variance of these powers in the past N [seconds] are calculated.

(A2) Frequency of Change in Voice Power

First, the logarithmic power of the recorded audio signal in 0.1 [seconds] intervals over the past N [seconds] at each time of self-report entry is calculated in the same way as (A1). Next, we count the number of intervals wherein the power exceeds threshold T_a in each interval during the past N [seconds].

4.3.2 Visual Features

(V1) Motion on the Table

Based on our finding that eating and drinking trigger relatively high levels of human interruptibility, the changes of image intensity inside the table area (*hotspot R*) were extracted based on the frame difference technique.

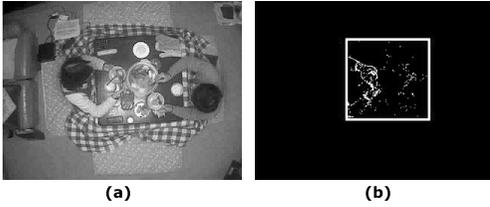First, difference image I' (previous frame and

Fig. 6 Results of motion tracking on the table. (a) shows input image. (b) shows the extraction results. White rectangle shows table area. White pixels show the changes due to motion. Note that the table area was manually extracted.

Fig. 7 Visual tracking results. (a) shows an input image. (b) shows updated background image. (c) shows occupants' location. White areas show occupant's location. White rectangle shows the table area. White circle shows circle enclosing occupants' mass. Note that the table area was manually extracted.

current frame) at time t of self-report entry is calculated as follows.

$$I'(x,y,t) = |I(x,y,t) - I(x,y,t-1)| \quad (1)$$

To suppress the noise created by changes in illumination, every pixel whose frame difference (absolute value) was smaller than threshold T_V was set to zero as follows. **Figure 6** shows the results of motion tracking. In this figure, each white pixel represents a change due to motion.

$$I''(x,y,t) = \begin{cases} 0 & (I'(x,y,t) < T_V) \\ I'(x,y,t) & (I'(x,y,t) \geq T_V) \end{cases} \quad (2)$$

Next, the sum of $I''(x,y,t)$ inside R is calculated.

$$Sum(t) = \sum_{(x,y)\in R} I''(x,y,t) \quad (3)$$

Finally, the average and variance of $Sum(t)$ over the past N [seconds] from time t is used to more stably extract these features.

(V2) Motion on Areas other than the Table

Based on our finding that the static behavior sleeping triggers low levels of interruption, features like occupants' motion are extracted. These features are calculated in the same way as (V1).

(V3) Relation in Position among Occupants

From our observations, we hypothesized that occupants are close to each other when engaged in the same behavior (eating and talking etc.) and cooperative work (teaching etc.), which suggests low levels of interruptibility.

First, occupants' locations were identified by using a background subtraction technique and background images. To robustly cope with changes in the environment, the background image $B(t)$ [13] is updated as follows. Note that image processing operations like erosion and dilation [11] are performed intermittently to sup-

press the noise caused by illumination and sunlight etc.

$$B(t) = k * I(t) + (1 - k) * B(t-1) \quad (4)$$

k is the leaning.

Next, the radius of the smallest circle that encloses occupants' masses at time t is calculated. Finally, average and variance of the radius over the past N [seconds] from time t are used as the features.

Figure 7 shows the results. In Fig. 7 (c), the white pixels show occupants' masses, and the white circle shows the circle enclosing the occupants.

(V4) Relative Change in Location

Based on our analysis, we found that static behavior like sleeping yields the lowest level of human interruptibility.

First, occupant's mass is extracted by the same way as (V3). The center of occupants' mass $(m_x(t), m_y(t))$ is then calculated. Furthermore, the relative change, $MR(t)$, in the center of occupant mass $(m_x(t), m_y(t))$ is extracted as follows.

$$\begin{aligned} MR(t) = &|m_x(t) - m_x(t-1)| \\ &+ |m_y(t) - m_y(t-1)| \end{aligned} \quad (5)$$

Finally, the average and variance of the relative change $MR(t)$ over the past N [seconds] from time t are used as the features.

4.4 Interruptibility Estimation

The level of human interruptibility was estimated using a trained Support Vector Regression (SVR) system. SVR is an optimization-based approach for solving machine learning regression problems based on the Support Vector Machine (SVM), which is widely used for human modeling and pattern recognition [14]. The system was constructed using a widely available, open source library (LIBSVM) [15].

5. Experimental Results

Several experiments were conducted to eval-

uate the effectiveness of our method. We used the same data (the 397 self-entered data points) mentioned in Section 3. 50% of the 397 self-entered data points were randomly chosen and used as training data. The rest (50%) were used as test data. We investigated how well our method could estimate the level of human interruptibility in the test data. The values of the level of human interruptibility estimated by our method were compared to the actual values given by the subjects.

First, to investigate the appropriate N (interval for calculating features) used in the proposed method, we calculated the correlation coefficients between estimated values and actual values at three intervals (5 minutes, 1 minute, 1 second) prior to the sample time. **Table 2** shows the results. **Figure 8** shows the distribution of actual values and those estimated by the proposed method. In this figure, the black line shows the straight line of regression. The interval of 1 minute yielded the highest value of correlation coefficient. Based on the results, below, N was set to 1 minute in the proposed method.

Second, to statistically investigate the significance of the correlation coefficient between actual and estimated values, the test of correlation coefficient ($\alpha = .01$) was applied. The results showed that there was a significant correlation between the two variables ($p < .001$). Furthermore, to evaluate the accuracy of the proposed method, we calculated R^2 (squared correlation coefficient) and $RMSE$ (root mean square error). **Table 3** shows the results.

As a practical standard, Onose[20] said that if the value of correlation coefficient ranges from 0.4 to 0.7, the correlation is pretty good. Hence, although the correlation of the proposed method is not so high, there is reasonable connection between the levels of human interruptibility (determined by subjective evaluation) and those estimated by the proposed method.

Next, we analyzed the cases of failure to estimate the level of human interruptibility and found that there were several causes as follows:
(1) Rapid changes in lighting (such as when an occupant switched on the lights) and sunlight degraded the accuracy of tracking the occupants' location due to the use of the background subtraction technique.
(2) It was difficult to distinguish something (such a large bag) that occupants placed in the room from the occupants. This degraded the

Table 2 Correlation coefficients for several intervals.

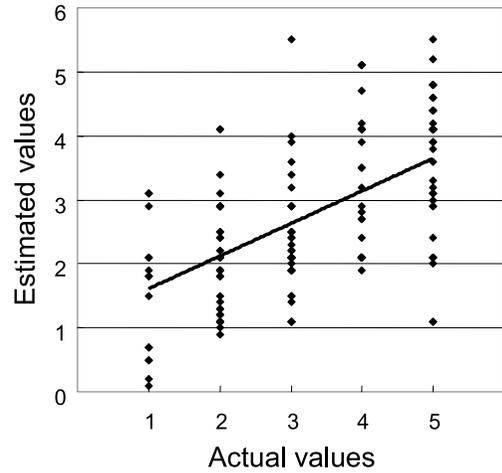| 5 minutes	1 minutes	1 second
0.53	0.57	0.49

Fig. 8 Distribution of actual human interruptibility scores and those estimated by the proposed method. The black line shows the straight line of regression.

Table 3 Results of accuracy. R^2 shows squared correlation coefficient. $RMSE$ shows root mean square error.

R^2	RMSE
0.32	1.02

accuracy of tracking the occupants' location.
(3) If the sound of the TV was large, the directional microphone captured the TV audio as part of the audio-data. This audio was incorrectly recognized as the voice of the occupant.
(4) Multiple occupants do not interact with each other. One subject engaged in an action (such as sleeping) that differed from that of the other occupant (writing on the table).

To quantitatively investigate the performance degradation yielded by the above errors, we eliminated (manually) 168 samples that evidenced the above errors and recalculated the correlation coefficient between actual and estimated human interruptibility. The recalculated correlation coefficient rose to 0.67. Hence, by developing techniques such as image processing and pattern recognition that can suppress the above errors, we can expect to improve the accuracy of the proposed method to some extent. Another source of error is that the same behavior may not be associated with the same level of interruptibility according to the time of day. Factors such as fatigue and physical condition

may alter the level of interruptibility.

6.　General Discussion

This section shows some challenges and future works for extending our method to create a really practical system.

6.1　Using the Results of Estimating Human Interruptibility

Although the accuracy of our current method is not so high, it is sufficient to develop practical applications according to how its results are used. We will continue to develop the following applications for smooth remote communication.

Since the caller wants to know the level of human interruptibility of the partner before calling, this information must be clearly and intuitively be conveyed to the caller without destroying the partner's privacy.

In the field of Computer Supported Cooperative Work (CSCW), for making remote communication smoother, many researchers have recently described methods that represent and convey *awareness information* to remote people. Itoh, et al. proposed a method in which optical fibers set on a potted plant gleam and rotate to indicate human presence and motion [16]. Takeuchi, et al. proposed a method that uses a door metaphor to represent the user's level of deskwork; a closed door means that the user is now busy [17]. The Digital Family Portraits used in the *Aware Home* change the size of a butterfly displayed on the portrait according to daily life activities such as opening and closing doors [18]. These works indicate that it is critical to intuitively and clearly visualize such information to enhance the attractiveness of remote communication.

Based on those ideas, we will design an intuitive visual representation of interruptibility level such as color. The color changes to show the level; red means a low level of interruptibily (don't call), yellow color (caution), and green means a high level of interruptibily (acceptable). Furthermore, we will also construct a criterion to evaluate the effectiveness of such visual representation and conduct experiments to investigate the improved smoothness achieved in starting communication.

Based on the estimation results, the volume of the telephone and answering machine mode are automatically set. For example, when the partner will be irritated by the interruption, the telephone is significantly or completely muted. This prevents the sound of the call from disturbing the partner.

6.2　Camera Position

In these experiments the camera was placed on the ceiling of the living room. This position has the merit of easily capturing a whole view of the living room. However, from the viewpoint of setting cost, this position is certainly not the most suitable. In the future, we must consider more practical camera positions for general amateur users. Furthermore, we will extract invariable features, i.e., those that are independent of camera position for other home environments.

6.3　Predicting Human Interruptibility

This paper focused on estimating current human interruptibility. An interesting challenge is predict human interruptibility. To this end, we will investigate which behaviors are good indicators of future levels of human interruptibility.

6.4　Improving the Estimation Accuracy

In the future, to improve the estimation accuracy, we will identify people in order to discriminate the intended partner from the others in the room. We will also investigate reasonable audio-visual sequential features (other than static features) to discriminate different behaviors that yield similar motions. We then will then attempt to discriminate them by using other context information such as day of the week and time of day etc. Furthermore, there are some interesting challenges with approaches that can directly extract human behavior such as sleeping, relaxing, which are closely related to human interruptibility as shown by our observation results.

6.5　Experiments Design for Capturing Human Interruptibility

This experiment is similar to pager research which uses the Experience Sampling Method (ESM) [19]. In the future, to capture the natural behavior of subjects, we must poll the subjects at irregular intervals instead of the regular 15 minute period used above. In this experiment, even if multiple occupants are present in the living room, the system output just one level of interruptibility. In the future, we will set multiple input devices that can capture each individuals' interruptibility, and then analyze the interruptibility of the individuals and the factors and related human behaviors in detail.

6.6　Extension to Rooms other than the Living Room

This paper focused on the living room. The

final goal of this research is to develop a method that can automatically estimate interruptibility of the communication partner wherever she/he is in the home. In the future, we intend to verify the proposed method in other rooms likes the entrance hall and kitchen. This will require us to set more sensors like cameras and microphones.

7. Conclusions

This paper introduced a method for automatically estimating and representing human interruptibility in a living room. To make online remote communication smoother, determining if it is appropriate to interrupt the remote partner is critical. Unlike most conventional studies on human interruptibility, this paper focuses on the home environment. The proposed method extracts several audio-visual features like body motion and voice from the data streams collected by a camera and a microphone. Based on these features, the level of interruptibility is estimated by applying the trained Support Vector Regression (SVR) technique. We conducted evaluation experiments to verify the accuracy with which the level of interruptibility could be estimated by the proposed method. The results show that the proposed method outputs reasonable agreement as to the level of human interruptibility. Furthermore, the method is practical since it is simple to implement, it uses only one cheap USB camera, a microphone, and one PC. Finally, we show the some challenges and future work for extending the proposal to yield a really practical system.

References

1) Rowan, J. and Mynatt, E.D.: Digital Family Portraits Field Trial: Supporting for Aging in Place, *Proc. CHI '05*, pp.521–530 (2005).
2) Mcfarlane, D.C.: Coordinating the Interruption of People in Human-Computer Interaction, *Proc.INTERACT '99*, pp.295–303 (1999).
3) Hudson, S.E., Fogarty, J., Atkeson, C.G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J.C. and Yang, J.: Predicting Human Interruptibility with sensors: A Wizard of Oz Feasibility Study, *Proc. CHI '03*, pp.257–264 (2003).
4) Minakuchi, M., Takeuchi, T., Kuramoto, I., Shibuya, Yu. and Tsujino, Y.: An Automatic Estimation Method for Busyness at Deskwork, *J. Human Interface Society*, Vol.6, No.1, pp.69–74 (2004) [In Japanese].
5) Horvitz, E. and Apacible, J.: Learning and Reasoning about Interruption, *Proc. ICMI '03*, pp.20–27 (2003).
6) Avrahami, D. and Hudson, S.: Responsiveness in Instant Messaging: Predictive Models Supporting Inter-Personal Communication, *Proc. CHI '06*, pp.731–740 (2006).
7) Fogarty, J., Hudson, S.E. and Lai, J.: Examining the Robustness of Sensor-based Statistical Models of Human Interruptibility, *Proc. CHI '04*, pp.207–214 (2004).
8) Fogarty, J., Ko, A.J., Aung, H.H., Golden, E., Tang, K.P. and Hudson, S.E.: Examining Task Engagement in Sensor-based Statistical Models of Human Interruptibility, *Proc. CHI '05*, pp.331–340 (2005).
9) Kidd, C.D., Orr, R., Abowd, G.D., Atkeson, C.G., Essa, I.A., MacIntyre, B., Mynatt, E., Starner, T.E. and Newstetter, W.: The Aware Home: A Living Laboratory for Ubiquitous Computing Research, *COBUILD '99*, (1999).
10) Yamazaki, T.: Ubiquitous Home: Real-Life Tested for Home Context-Aware Service, *Proc. TRIDENTCOM '05*, pp.54–59 (2005).
11) Okutomi, M. (Eds.): *Digital Image Processing*, The Computer Graphic Arts Society (2004).
12) Krumm, J., Harris, S., Barry, M., Brumitt, B., Hale, M. and Shafer, S.: Multi-Camera Multi-Person Tracking for EasyLiving, *VS '00*, p.3 (2000).
13) Lipton, A., Fujiyoshi, H. and Patil, R.S.: Moving Target Detection and Classification from Real-time Video, *Proc. 1998 Workshop on Applications of Computer Visions*, pp.8–14 (1998).
14) Burges, C.J.C.: A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery*, Vol.2, pp.121–167 (1998).
15) http://www.csie.ntu.edu.tw/˜cjlin/libsvm/
16) Itoh, Y., Miyajima, A. and Watanabe, T.: 'TSUNAGARI' Communication: Fostering a Feeling of Connection between Family Members, *Ext. Abstracts CHI '02*, pp.810–811 (2002).
17) Takeuchi, T., Shibuya, Y. and Tsujino, Y.: Detection and Presentation of Awareness Infor-

mation for Remote Communication, *Proc. Human Interface Symposium*, pp.581–584 (2001) [In Japanese].

18) Mynatt, D.E., Rowan, J., Jacobs, A. and Craighill, S.: Digital Family Portraits: Supporting Peace of Mind for Extended Family Members, *Proc. CHI '01*, pp.333–340 (2001).

19) Kubey, R.W. and Czikszentmihalyi, M.: Television and the Quality of Life: How Viewing Shapes Everyday Experience. Lawrence Erlbaum Assoc Inc. (1990).

20) Onose, S.: Analysis of Statistical Data, Uchida Rokakuho Publishing Co., Ltd. (1995).

Yoshinao Takemae received B.E., M.E. and Ph.D. degrees from Keio University, Japan, in 1999, 2001 and 2006, respectively. In 2001, he joined NTT Corporation. From 2001–2004, he worked in NTT Communication Science Laboratories. Since 2005 he has been a researcher at NTT Cyber Solutions Laboratories. His research interests include human tracking and supporting remote communication. He is also a member of IEICE and ACM.

Takehiko Ohno received B.E. and M.E. degrees from Tokyo Institute of Technology, Japan, in 1992 and 1994, respectively. In 1994, he joined NTT Corporation. His research interests include gaze-based interaction, gaze tracking technologies, computer-mediated communication, information appliances, and cognitive models of human-computer interaction. Now he is a manager of Information and Communication Technology Office at Department 2 in NTT Corporation. He is also a member of ACM and JCSS.

Ikuo Yoda received the B.E. and M.E. degrees in Electrical Engineering from Waseda University, Tokyo, Japan in 1986 and 1988, respectively. In 1988, he joined NTT where he engaged in R & D of TMN-based network management systems. He is now a Senior Manager at NTT Cyber Solutions Laboratories. His research interests are the home network and service management. He is also a member of IEICE.

Shinji Ozawa received B.E., M.E. and Ph.D. degrees all in Electric Engineering from Keio University, Japan, in 1967, 1969 and 1972, respectively. He was a research associate, Assistant Professor and Associate Professor starting in 1970, in the department of Electric Engineering of Keio University. He is now a Professor of the Department of Information and Computer Science. In 1984 he was a Visiting Processor at the University of Maryland, U.S.A. He has worked on Digital Communication and signal processing for images and speech. He is a Fellow of IAPR. He is also a member of IEICE and IEEE.