Journal of Mobile Multimedia, Vol. 8, No. 2 (2012) 132–148

IMPLICIT CONTEXT AWARENESS BY FACE RECOGNITION

SEIJI TAKEDA
Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho
Nada-ku, Kobe, Hyogo, 657-8501, Japan
takeda-1@stu.kobe-u.ac.jp

TSUTOMU TERADA
Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho
Nada-ku, Kobe, Hyogo, 657-8501, Japan
tsutomu@eedept.kobe-u.ac.jp

MASAHIKO TSUKAMOTO
Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho
Nada-ku, Kobe, Hyogo, 657-8501, Japan
tuka@kobe-u.ac.jp

Received February 1, 2012
Revised April 26, 2012

In recent years, technical improvements to sensors have attracted a great deal of attention, in particular due to the sensors' capability recognizing user contexts. In this paper, we propose an implicit context awareness system that identifies user context by sensing the context of surrounding environments. We implemented a prototype that recognizes user contexts by sensing surrounding people by two cameras We actually used the prototype in a variety of situations. Evaluation results showed that the system was effective and improved context recognition. Our method can be used to identify rich contexts that cannot be recognized by conventional methods.

Keywords: an implicit context awareness, face recognition, wearable computing

Communicated by: D. Taniar & E. Pardede

1 Introduction

The recent downsizing of computers has attracted a great deal of attention to the area of wearable computing. In wearable computing environments, computers support the users anytime and anywhere. Needless to say, there are many potential applications for daily living. The usability of these application services is on the base of user's situations or environment as they walk, sit, talk while walking, or simply spend time in a crowd. Therefore researchers have recently been focusing on context awareness and its role in various wearable sensors, such as accelerometers, global positioning systems (GPS), camera, and infrared sensors. Wearable computers provide users with optimized services by recognizing user contexts: for example, by displaying an alert to buy milk when the user is near a market or by displaying a train timetable when the user enters a train station. With such services, it is important to recognize user context with high accuracy and stability. There has been much research on ways to improve the recognition accuracy and the recognition of new contexts by using various sensors[5, 6, 9]. In these studies, user context is recognized directly via wearable sensors such as accelerometers or biological sensors. Although such systems can identify user motions and health conditions in detail, they cannot recognize contexts that depend on surrounding environments, such as being watched by surrounding people, being in crowded place, eye fatigue due to illumination and so on.

In this paper, we propose an implicit context awareness system that recognizes user contexts by sensing the context of the surrounding environments with a wearable camera. These user contexts are recognized by translating the surrounding environment data. Implicit context awareness enables us to recognize user context from surrounding environments, which cannot be done with conventional methods sensing user's motion directly.

We assume that the people surrounding the user are the most important element of the surrounding environment. Because we are affected by surrounding people in various situation in daily life. For example, talking, walking and doing a meeting. Therefore our system detects the direction of these people's faces because this positioning indicates what they are paying attention to. For example, if most people are facing the user, it means that the user is attracting a lot of attention, perhaps because there is something attractive and/or strange about them. In such a situation, the user is under the context of *being watched* and the system alerts him to check his appearance because there is a possibility that, for example his fly is open. In our prototype, the system recognizes detailed contexts by the integrative use of conventional context recognition methods and our proposed implicit context recognition system. When the system recognizes user's situations, for example the user is being watched by implicit context awareness, it determines the reason by combining implicit and explicit contexts.

This paper is organized as follows. In Section 2 we discuss related work, and in Section 3 presents the proposed method. Section 4 explains the implementation of our method and experiments are discussed in Section 5. Finally, we conclude the paper in Section 6.

2 Related work

In recent years, technical improvements to sensors have attracted a great deal of attention, in particular due to the sensors' capability recognizing user contexts. Therefore, contexts awareness is used in wide field like swimming[2], navigation[3].

In the Bachlin's study[2], user wear three accelerometers and three actuator that LED in goggles1, vibration swim pants and audio in cap. Then user is supported by actuator based on the accelerometers data. For example, if user swims bad form, LED lights or vibration vibrate. Therefore user is able to swim well in short time practice. This study shows wearable computing been able to use in water.

As an example of improvement of recognition rate, Nguyen Dang Binh proposed a technique that improving person detection by camera[4]. In this study, he made cooperative learning framework that is generate by generative model that is trained robustly on motion data and discriminative model that is labels training data. Then the system can recognize person over 90% rate in train station and in the school and so on.

As an example of combining a variety of sensors to recognize detailed contexts, Aoki, et. al. proposed a technique of context awareness that uses not only information from electrical equipment's ON/OFF state or infrared sensor but also from an omnidirectional camera that recognizes the position and posture of people in a room[5]. Using this camera enables the system to accurately recognize detailed contexts. There has also been research on a system that recognizes a user's position and direction via a wearable camera that captures surrounding images by using a relative angle sensor,

Fig. 1. Goggles with one bi-color LED[2]

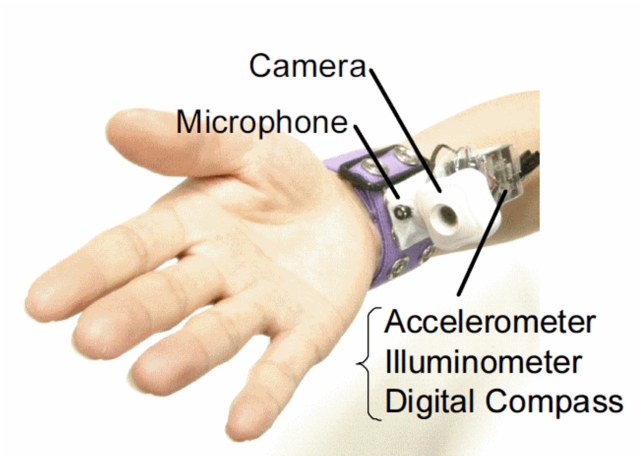

Fig. 2. Device of Maekawa's study[7]

a magnetic sensor, and an accelerometer[6]. In addition Maekawa, et. al. proposed a technique of context awareness that uses camera, microphone, accelerometer, illuminometer and digital compass wearing wrist[7]. Figure 2 shows prototype device. In this study, wearing sensors on wrist, so system can recognize detail of human action do with the hands.

On the other hands, there are some methods that recognize detailed contexts by using low number of sensors like accelerometer only or camera only. Blanke, et. al. proposed a new method to recognize user's position in office or in the house by using only 3D-accelerometer and gyroscope[8]. In this study, they used only the data from 3D-accelerometer and gyroscope however the system recognize user position over 70% that is very high.

A tool has been developed using the recognition technology mentioned in the above researches[9][10]. In Okuma's study[9], it performs a guide service for an exhibition in the museum3. In this study, an experiment was performed in a science museum in Tokyo's Kitanomaru Park. The system detects user

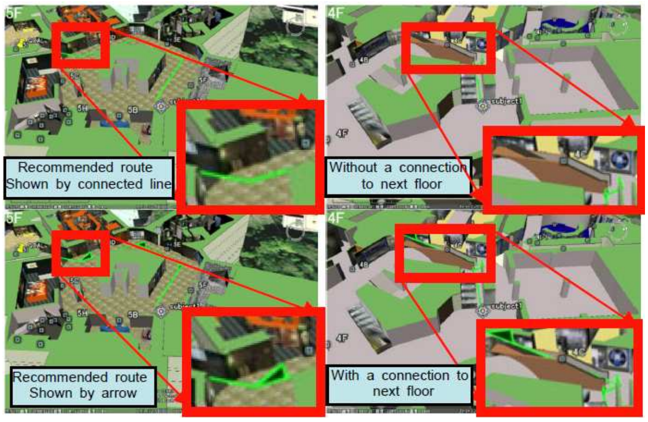

Fig. 3. Example of induction content[9]

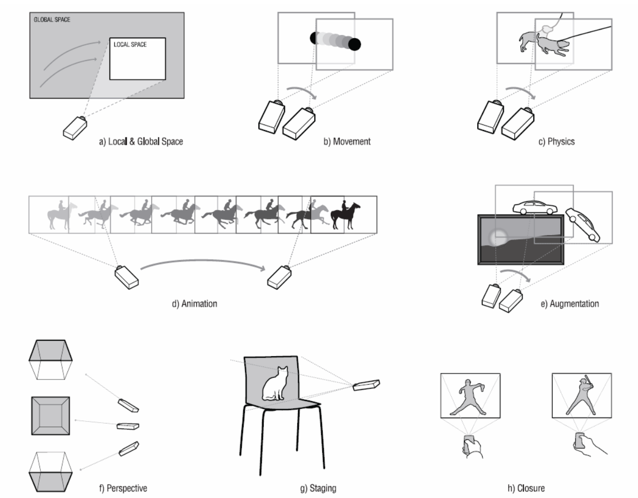

Fig. 4. MotionBeam interaction principles[10]

position and posture with a magnetic sensor, gyro sensor, accelerometer sensor, and RF-ID tag. Wi-Fi access points and RF-ID readers were placed in specific points throughout the museum. The system showed user position, description on exhibits. In Wills's study[10], author made some applications using user's motion. In this study, user can control the character or movie by device because devices recognize user's hand motion4.

Almost of conventional research in the field of recognizing context awareness has not adequately considered surrounding environments but rather focused only on data that sensing user motion directly.

There is a conventional research that considering surrounding environment. [11] Nakamura, et. al. using not only an acceleration sensor but also a carbon dioxide sensor, which is worn on the hip to recognize user contexts. This sensor recognizes limited user contexts by analyzing the ambient environment concentrations of carbon dioxide. However, there is a problem that if only carbon dioxide is used for the ambient situations. This study can use it only in limited circumstances such as in a room

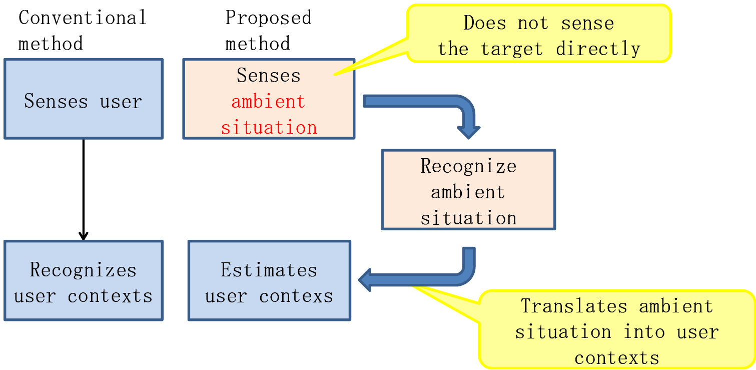

Fig. 5. Comparison of proposed and conventional methods

or on a train. Moreover, system cannot recognize many user contexts via only carbon dioxide sensor.

3 Proposed Method

3.1 Implicit context awareness

Conventional methods recognize user contexts directly through wearing wearable accelerometers or biological sensors or capturing a camera to recognize user motions. These methods are using one sensor data directly or using various sensor data with combining to recognize user context.

In contrast, in implicit context awareness, the camera, thermography, a thermometer, GPS, infrared sensors, and ultrasonic sensors are used not for sensing the target user, however they are used for sensing the target's surrounding environments. This type of sensing, provides us with environmental data related to the target, such as people's information that may be relevant to user's condition, surrounding sound, temperature, and humidity. These data can be then translated into user context data, thus enabling us to recognize user context indirectly. This system can recognize the condition based on the surrounding environment. For example, moving in crowd place or been focusing that it is difficult to recognize conditions in the conventional method. The proposed method senses the surrounding situations to recognize how user contexts are affected by them, this enables us to recognize situations such as whether or not a user is being watched by surrounding people, the appropriate walking speed considering the surrounding people, or the appropriate speaking volume. These recognitions expand the range of context recognition.

The differences between the conventional and proposed methods are shown in Fig 5. The proposed method differs in two primary ways:

- It does not sense the target directly.
- It senses surrounding situations related to the target, then translates this data into user context.

3.2 System design

There are two requirements in implicit context awareness. First, the recognition of surrounding conditions concerned with user conditions is needed because the user condition is determined by the surrounding environments. Second, it must be possible to recognize user context by using data about

Fig. 6. Examples of situations that can be estimated from ambient environments

Sensors	Recognizing ambient situation	User contexts
Camera	Ambient environmental changes	There is something you should focus on
	Human walking speed	Appropriate walking speed
	Detects many front faces	Has attracted attention
Thermography	Ambient temperature	Whether user is indoors or outdoors
	Ambient temperature change	Whether wearing a jacket or not is better
Hygrometer	Ambient humidity	Weather
Infrared sensors	Whether other people are around	User may speak loudly
	Distance between people around user	Degree of congestion
Microphone	Volume of ambient sound	Volume of speaking
	Type of avient sound	Location estimation

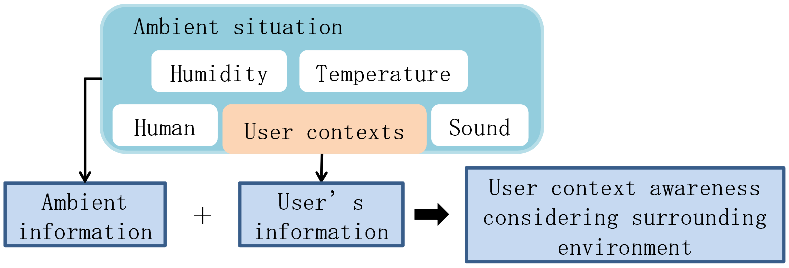

Fig. 7. Flowchart

the surrounding environments. In order to clarify these requirements, Table 6 shows examples of sensors that can be used for implicit context awareness.

As seen in the table, a variety of situations can be recognized because there are many kinds of sensors, moreover, each sensor can recognize different surrounding environments. It is clear that the system can recognize user contexts that have previously been unrecognized because it considers the surrounding environments.

Moreover, if we combine user contexts from the proposed method with conventional method that recognized user contexts directly, we can recognize user contexts in more detail. For example system can recognize user should turn light because study in dark place or user should not do boring presentation and so on.

Figure 7 shows the flowchart of the proposed method's processes. Each process is shown in detail at the following sections.

3.3 An implicit context awareness by face recognition

As mentioned previously, the direction of surrounding people's faces has an influence on user behavior. Moreover there are many techniques for face recognition[12][13]. We therefore use information on face direction to look for implicit context awareness.

3.3.1 Obtaining information about the surrounding situation

The procedure for obtaining information about the surrounding circumstances is outlined below. Figure 8 shows a screen shot of this procedure that was captured by two cameras the user wore on the front and back of the body. Figure 9 shows a rough flowchart of the procedure, which is outlined

Fig. 8. Screen shot of image acquisition

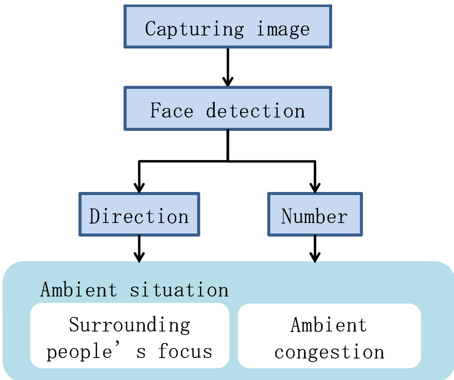

Fig. 9. Flowchart of recognizing user context

below.

1. Capturing surrounding situation by camera.

   The system captures surrounding images in real time by using an every direction camera or multiple cameras.

2. Detection of direction of the face

   To detect the face, system has to calculate feature quantity. In this system calculate feature quantity by adapting various pattern Haar-like features[14]10 in various positions. Feature quantity is calculated by this formula1.

$$featurequantity = totalpixeloftheblack - totalpixelofthewhite \quad (1)$$

   After calculating feature quantity, system detects the feature positions if the feature quantity that is calculates is near to the feature quantity that is previously learned. In this way, it seems that system takes long time. However system ignore the pixel that is no object and connecting the many kinds of detectors in strength of detection order. Therefore system can detect feature points very speedy. In this research, system detects what surrounding people are looking at by

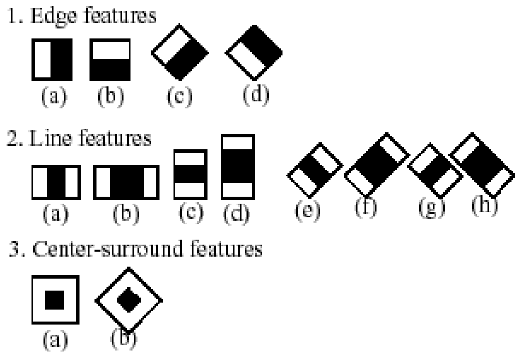

Fig. 10. Haar-like features

Quote from (http://opencv.willowgarage.com/documentation/object_detection.html)

detecting the face direction (right, front, or left). This processing system that detects face is provided by OpenCV[15].

In this paper, we define the name of face direction based on people who sensed. Therefore if people who sensed is looking left side of him, we define the face direction's name is "left". We do the same for front and right.

3. Translate to surrounding situations

   We use cameras to obtain data on the surrounding environment, number of front, right, and left faces and the relative location between a user and the surrounding people. We can recognize the surrounding situations by the number of faces pointed in each direction and by the detected face's location. In the proposed method, we can use two elements. First element is the type of face detected: for example, front face is only detected, right face is only detected, or left and front faces are detected. Second element is the number of detected faces. Transition of the number of detected faces is also useful element. Table 11 shows lists of the surrounding circumstances and conditions used to estimate the situation around the user by using face information. The example is of a user who is wearing two cameras, one on the front of the body and one on the back. (-) means items were not involved in the recognition result. The face counts in Table 11 are decided tentatively by preliminary study.

Using implicit context awareness enables us to recognize situations that were not previously recognized. In addition, if we use this method combined with the method that recognizes contexts directly, the system can recognize more detailed contexts like those in Table 12.

3.4 System using implicit context awareness

It is possible to create new applications by using an implicit context awareness method. One example is an application that creates an awareness of the lack of strange shapes or of user behavior by sounding an alarm when the system recognizes that a user is being watched by surrounding people because many front faces are detected. Moreover, using other sensors, we can make applications that have

Fig. 11. Example of translating information about the face to ambient situations

Direction of face	Camera that detects	Condition	Ambient situation
-	-	Total is more than five	Ambient situation is crowded
-	-	Total is less than five	Ambient situation is quiet
Front face	-	Total is more than five	User is watched by surrounding people
Left face	Front side	Sum of front and back sidesis is more than five	Surrounding people are looking at the user's right side
Right face	Back side		
Right face	Front side	Sum of front and back sides is more than five	Surrounding people are looking atthe user's left side
Left face	Back side		
Front face only	Only front side	Total is more than five	People around user walk a certain direction
	Only back side		
-	-	Face detection has stopped	People around user begin to move
-	-	No face detect	There are no people in the situation

Fig. 12. Example of combination of circumstances surrounding environments and self-recognition results

Recognition result	Own situation	Recognition by implicit context awareness
Being watched by surrounding people	Do nothing	Dress or costume is odd
	Giving a presentation	Giving a good presentation
	Walking	Walking while receiving the attention
Surrounding people look a certain way	Walking in the city	Something interesting is happening nearby
Being crowded	Having a conversation	Should be considering volume of voice
	Using mobile phone	Should stop talking on mobile phone
	Walking	Should walk carefully because it is crowded
Moving ambient people	Waiting signal	Should walk because signal is changing to green
Sudden attention	Having a conversation	Remarking on something strange
	Using mobile phone	Surrounding people feel bad about talking on mobile phone

various functions. If the user wears a geomagnetic sensor on the head, we can make an application that sounds an alert to make user look in the direction most of the surrounding people are looking. In addition, if user is using GPS, we can make an application that creates a map based on a photographic image of what many surrounding people paid attention. To realize these applications, we need a framework that enables us to recognize situations like many surrounding people looking in the same direction. We also need the platform to manage and to integrate a variety of sensors. In this study, we used *Wearable Toolkit*[16, 17, 18], which is a sensor management and service delivery platform for wearable computing.

3.4.1 Wearable Toolkit

Wearable Toolkit, a system platform for wearable computing, is a middleware that runs between OS and wearable services. In Wearable Toolkit, we describe system behaviors with event driven rules and use plug-ins to extend system functions. Service developers can easily construct various services by combining plug-ins[19] and describing rules that use the functions of plug-ins. We can add, delete, or

```
DEFINE    Rule-ID
      [ IN   List-of-belonging-groups ]
      [ FOR   Scope ]
      [ VAR   Variable-name   AS   Variable-type ]
WHEN   Event-type [ ( Target-of-event ) ]
IF        Conditions
THEN DO   Actions
```

Fig. 13. Syntax of ECA rules in Wearable Toolkit

Fig. 15. NEW.DATA corresponding to type of information

Name	Type of data
NEW.DATA1	Number of front faces by front camera
NEW.DATA2	Number of front faces by back camera
NEW.DATA3	Number of front faces by both cameras
NEW.DATA4	Number of side faces by front camera
NEW.DATA5	Number of side faces by back camera
NEW.DATA6	Number of side faces by both cameras
NEW.DATA7	Number of faces by both cameras

customize those services while the system is running by adding, deleting, or modifying the rules.

All services are represented as a set of ECA rules, which were originally a behavior description language in an active database system that is a database technology. An active database processes the prescribed actions in response to an event arising inside or outside the database.

3.4.2 ECA rule

Each ECA rule consists of three parts: Event, Condition, and Action. The event part describes an event that occurs in the system, the condition part describes the conditions for executing actions, and the action part describes the operations to be carried out. Figure 13 shows the syntax of ECA rules for Wearable Toolkit. In the Figure, *Rule-ID* describes the name of the ECA rule, and *Event-type* describes the name of the event that triggers this rule. *Conditions* specifies the conditions for executing the following actions, and we can use AND and OR operators in *Conditions* for describing complicated conditions. *Actions* specifies executing operations and their arguments.

3.5 Application example

These items are example of situation that user's behavior affected by surrounding people's faces in daily life.

- Conversation with person
- Waiting for a signal
- Watching the same thing or looking in the same direction
- During a presentation
- Being crowded place

```
//Initialization
DEFINE Startup
WHEN CMN_SYSTEM_START
THEN DO SERIAL_INIT_FORMED('COM2', 100, 9600, 8, 0, 0, 0,
'¥ [ˆ¥ n]+¥ ,¥ [ˆ¥ n]+¥ ,¥ [ˆ¥ n]+¥ ,¥ [ˆ¥ n]+¥ ,
¥ [ˆ¥ n]+¥ ,¥ [ˆ¥ n]+¥ ,¥ [ˆ¥ n]+¥ ¥ n', 1)
DO GLOBAL.CONTEXT = STRING('NONE')
// Start the context tool
DO START_CONTEXT_TOOL()

// 1.If the number of faces detected is above threshold,
// inform the user that ambient environment in busy.
DEFINE Crowded
WHEN SERIAL_FORMED_DATA IF NEW.DATA7 > '4' THEN
// Remember that ambient environment is crowded
DO GLOBAL.CROWDED = BOOL(1)
DO GLOBAL.EMPTY = BOOL(0)
DO CMN_ALERT('Environment is crowded')

// Remember that ambient environment is quiet (sum of detected faces is less than four)
DEFINE Few
WHEN SERIAL_FORMED_DATA IF NEW.DATA7 < '5' THEN
DO GLOBAL.CROWDED = BOOL(0)
DO GLOBAL.EMPTY = BOOL(1)
DO CMN_ALERT('Environment is quiet')

// Alert user that surrounding people are moving
// when suddenly no more faces are detected DEFINE Change
WHEN SERIAL_FORMED_DATA IF ?GLOBAL.INCOMING AND NEW.DATA7 < '5' THEN
DO GLOBAL.CROWDED = BOOL(0)
DO CMN_ALERT('Surrounding people are beginning to move')
```

Fig. 14. Examples of proposed applications (1)

- Using a mobile phone in a room or train

Conventional methods cannot recognize these items. But implicit context awareness can recognize these situations. So we can create new applications that are aware of the user's situation by using the proposed implicit context awareness method. Because of this point, it is appear that implicit context awareness is useful.

Below are some examples of services we have built.

1. The system informs the user that the surrounding environment is crowded when the number of detected faces exceeds the threshold.

2. Inform user that surrounding people have begun to move when detected faces are suddenly no longer detected.

3. Warn user, in either words or action, that there is something wrong when it detects a front face above the threshold while user does nothing.

4. Take pictures and inform user if there is an interesting thing when it detects a side face that is right or left above the threshold.

5. Let user decide whether or not to respond to a Skype call when it detects a face under the

```
// 3.Alert user to funny dress or behavior?
// when detect over threshold of front face.
DEFINE Focus
WHEN SERIAL_FORMED_DATA
IF NEW.DATA3 > ’4’ THEN
DO CMN_ALERT(’Being watched by surrounding people, is your dress or behavior funny?’)

// 4.Alert that there may be interesting thing around
// when detecting side face above threshold.
DEFINE Camera
WHEN SERIAL_FORMED_DATA
IF NEW.DATA6 > ’4’ THEN
DO SERIAL_SEND(’COM2’,’1’)
DO CMN_ALERT(’Surrounding people are looking at criterion. It anything happening?’)

// Using Skype
// 5.System lets user decide whether to respond to Skype call
// when detects face subthreshold, system refuses Skype’s incoming call
// when detects front face above the threshold.

// Remember that have not yet recalled for Skype’s incoming call.
DEFINE Incoming
WHEN SKYPE_INCOMING THEN
DO GLOBAL.INCOMING = BOOL(1)

// refuse Skype’s incoming call when detects front face above the threshold.
DEFINE Crowdedincoming
WHEN SKYPE_INCOMING
IF ?GLOBAL.INCOMING AND ?GLOBAL.CROWDED THEN
DO SKYPE_HANDLE_INCOMING(NEW.ID, ’FINISHED’)
DO GLOBAL.INCOMING = BOOL(0)
DO GLOBAL.REJECT = BOOL(1)
```

Fig. 16. Examples of proposed applications (2)

subthreshold. System also can refuse incoming Skype call when it detects a front face above the threshold.

6. Take a photo when it detects above the threshold and user performs a gesture that has been done before and learned.

Figures 14 and 16 show examples of the rules used to make the proposed applications. In these examples, the threshold number is five for the sake of simplicity. Table 15 shows which NEW.DATA corresponds to different of types of information sent via serial communication.

4 System implementation

We implemented a prototype of the proposal system. We used a Lenovo Thinkpad X200 computer (CPU: Intel(R) Core2Duo P8600 2.40 GHzCRAM: 3 GB) with a Windows XP operating system, Microsoft VisualC++2005 for application, WearableToolkit(rule engine dll ver. 0.1.0.69) , OpenCV(Open Source Computer Vision Library) that is a collection of computer vision library developed by Intel Corporation and videoInput[20] that is a library make easily to use the DirectShow video capture, com0com [21] ver. 2.2.2.0 is for tool that virtual ports. As a camera, we used two Microsoft LifeCam Show RLAs made by Microsoft-00007 (image sensor: two million pixels, focal length: 50 cm - 1.5 m, F-number: 2.0)

```
// Reply to encourage people to call denied before.
DEFINE Return
WHEN SERIAL_FORMED_DATA
IF ?GLOBAL.REJECT AND ?GLOBAL.EMPTY THEN
DO CMN_ALERT('Are you sure you do not want to reply to that call that you refused?')
DO GLOBAL.REJECT = BOOL(0)

// Fewer people around, alert when there is an incoming call.
DEFINE Incomingalert
WHEN SKYPE_INCOMING IF ?GLOBAL.INCOMING AND ?GLOBAL.EMPTY THEN
DO CMN_SHOW_QUESTION(NEW.ID, 'Called by NEW.DISPNAME. Would you like to receive?')

// If user answers YES, receive.
DEFINE SkypeAccept
WHEN CMN_ANSWER IF ?NEW.RESULT AND ?GLOBAL.INCOMING
THEN DO SKYPE_HANDLE_INCOMING(NEW.ID, 'INPROGRESS')
DO GLOBAL.INCOMING = BOOL(0)

// If user answers NO, refuse.
DEFINE SkypeReject
WHEN CMN_ANSWER IF !NEW.RESULT AND ?GLOBAL.INCOMING
THEN DO SKYPE_HANDLE_INCOMING(NEW.ID, 'FINISHED')
DO GLOBAL.INCOMING = BOOL(0)

// Take a picture when doing learned gesture
// under the situation detecting face above the threshold.
DEFINE Shutter
WHEN CONTEXT_RECOGNIZED(shutter) IF ?GLOBAL.CROWDED THEN
DO SERIAL_SEND('COM2','1')
```

Fig. 17. Examples of proposed applications (3)

Next, we describe the details of the implementation.

4.1 Image processing

In our implementation of the prototype, we used training data for the faces that had been prepared in OpenCV. We used only the side face, not left or right, because the prepared training data could not discern the right face from the left. Figures 18 and 19 each show an example of recognition of the front and side face, respectively, using the training data we uses in the prototype. In prototype, if system detect front face then system write circle on the front face and shows "front" under the face. Moreover if system detect side face then system write square on the side face and shows "side" under the face. According to Figure 20, system recognize front face and side face in one face in the case that people look slanted. In addition there are some mistaken that do not detect face however there are faces. Moreover show the example detect face in no face position in Figure 21. Because of these mistakes, it seems to take mistake in experiment. However there is no problem because rate of mistake is very low. These examples faces were detected from still images.

4.2 Estimating self context aware

In this prototype, the ambient environment is recognized as crowded when more than five face detected. The ambient environment is recognized as quiet when less than five faces are detected. When the sum of detected faces is over five, it is estimated that the user is being watched by surrounding people regardless of the number of side faces present, and if the sum of side faces detected is over five, it is estimated that the surrounding people are looking at a specific thing regardless of the number of

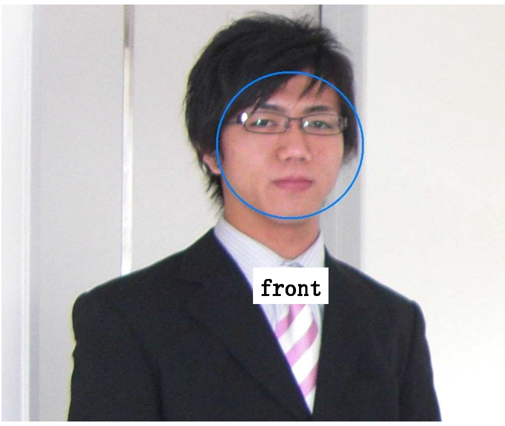

Fig. 18. Example of detection of front face

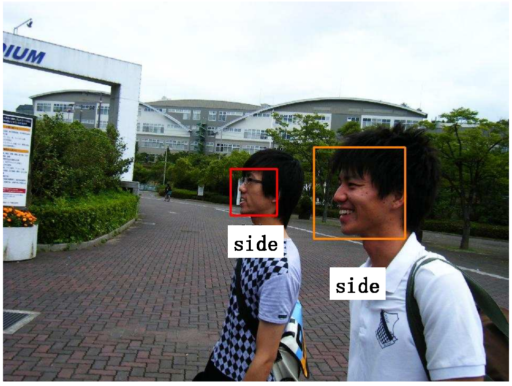

Fig. 19. Example of detection of side face

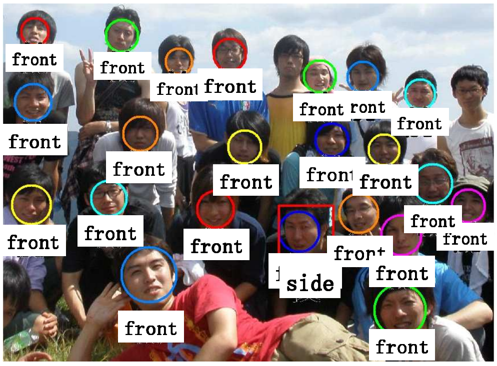

Fig. 20. Example of detection in group photo

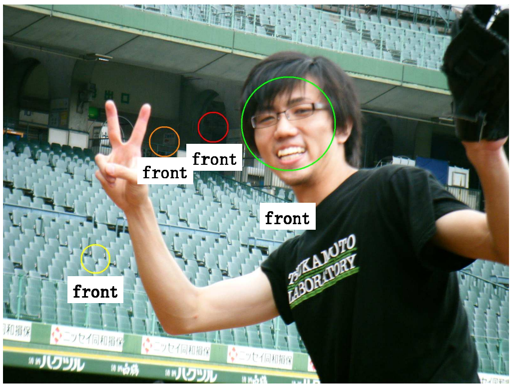

Fig. 21. Example of miss detection

front face detected. The data on estimate contexts is sent to Wearable Toolkit.

4.3 Connecting to Wearable Toolkit

The application that obtains the face data communicates with Wearable Toolkit via com0com, which is a virtual serial port in one computer. The number of faces pointing in each direction is sent to Wearable Toolkit.

5 Experiment

To determine the recognition rate accuracy of the implicit context awareness by face recognition system, we evaluate by checking match between the recognition results from using the prototype and from humans looking at video.

5.1 Preliminary experiment

For the preliminary experiment, we detected the congestion and the level of surrounding interest at Sannomiya arcade in Kobe and at Kobe University School of Engineering building in Kobe, Japan. Figure 23 shows the experiment set-up. In this experiment, a participant wore two Web cameras on his front and back to recognize surrounding situations. For labeling, we used eye movie GDV180 and

Fig. 22. Experimental results

User contexts	Label number	Recognition times	Missed recognition times	Precision	Recall
Surrounding area is crowded	9	7	2	77.8 (%)	77.8 (%)
Be in the spotlight	3	3	5	37.5 (%)	100 (%)

a video camera to record the front camera's and back camera's movie.

The camera's image size was small - 320 240 pixels - because if images from Web cameras are too big, our system cannot process them in real time.

The experimental results demonstrated that the system can recognize if the user is in a crowded place or being watched by surrounding people. However, these were some problems. The image size was too small, so the system could not recognize people far away from the user. Also, the system's frame rate was very low (5 FPS), which made it difficult to recognize people who were moving.

5.2 Evaluation experiment

The preliminary experiment demonstrates the it is difficult to recognize situations in real time. We therefore performed an additional experiment in which HD video was used for the evaluation. For getting HD video the participant walked around Sannomiya arcade with wearing compact digital cameras on the front and back of his body to obtain information on the surrounding situations. This experiment was carried out for 20 minutes.

The prototype performed face recognition on the video footage moreover recorded the number of detected faces hour by hour. We performed manual labeling while watching the parts of the movie when the user was in a crowd or being watched by surrounding people. The evaluation was performed by matching the results against recognition by the prototype. The threshold for recognizing "user being watched by surrounding people" or "ambient environment is crowded" was defined manually by looking at the first ten minutes of movie. Then evaluate by comparing with manual recognition and the system recognition using movie last 10 minutes.

5.2.1 Evaluation results

After watching the movie for ten minutes, we appointed three as the threshold for recognizing that a "user is being watched by surrounding people" moreover four as the threshold for recognizing that a "user is in a crowded place".

Table 22 shows the label number, recognition times, missed recognition times, precision, and recall. The proposed system's recall was 66.7% and its precision was 77.8% when the ambient environment was crowded, while the recall was 100% however the precision was 37.5% when the user was being watched by surrounding people. Unrecognized and missed recognition sometimes occurred, however the evaluation result shows that overall, implicit context awareness by face recognition is useful. The labels of users being watched are few because it is rude to stare at people in everyday life. We think that missed recognition of a user being watched is due to recognizing people's faces as they walk behind the user as "watching".

6 Conclusion

In this paper, we proposed an implicit context awareness method that recognizes user contexts by recognizing the surrounding situations, moreover discussed its implementation and design. Our sys-

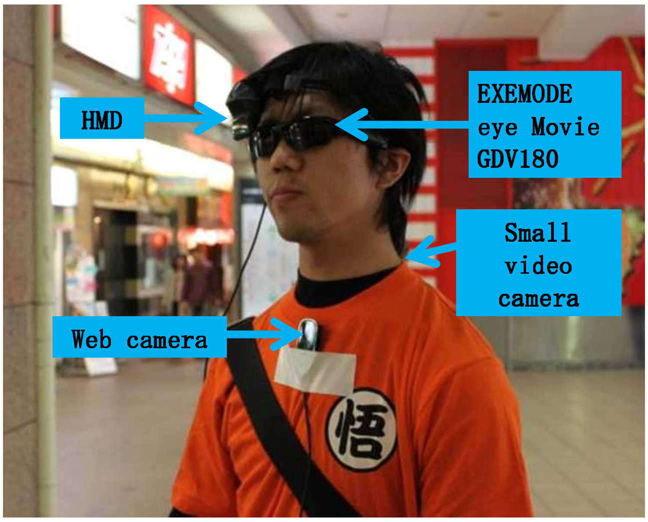

Fig. 23. State of the Experiment

tem can recognize new contexts that cannot be recognized by conventional methods because it was an implicit context awareness of the surrounding people's face direction. Result of prototype testing demonstrated that the recognition of surrounding situations is useful for recognizing user contexts. We feel that this information can be used to develop new tools or create new ways of using existing tools.

In the future, we aim to improve the system's processing speed and enable it to distinguish right faces from left. We also want to use not only face but also eye direction to improve the accuracy. Finally, we intend to use and combine additional surrounding information such as temperature, the movement of people, the degree of sunshine, and the type or amount of mail in a mail box for implicit context awareness.

Acknowledgments

This research was supported in part by a Grant in aid for Precursory Research for Embryonic Science and Technology (PRESTO) from the Japan Science and Technology Agency and by a Grant-in-Aid for Scientific Research(A)(20240009) from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

1. H. Okuda, T. Suzuki, A. Nakano, S. Inagaki and S. Hayakawa: Multi-hierarchical Modeling of Driving Behavior using Dynamics-based Mode Segmentation, *IEICE transaction on Fundamentals of Electronics, Communications and Computer Sciences (FECCS 2009)*, Vol. E92-A, No. 11, pp. 2763–2771, (Jan. 2009).
2. M. Bachlin, K. Forster and G. Troster: SwimMaster: A Wearable Assistant for Swimmer, *In Proc. of the 11th ACM international conference on Ubiquitous computing(Ubicomp 2009)*, pp. 215–224, (Sept. 2009).
3. H. Hile, R. Grzeszczuk, A. Liu, R. Vedantham, J. Borriello and G. Borriello: Landmark-Based Pedestrian Navigation with Enhanced Spatial Reasoning, *In Proc. of the 7th International Conference on Pervasive Computing (Pervasive 2009)*, pp. 59-76 (May. 2009).
4. Nguyen Dang Binh: Long Term Carefully Learning for Person Detection Application to Intelligent Surveil-

lance System, *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia(MoMM 2011)*, pp. 34–41, (Dec. 2011).

5. S. Aoki, Y. Iwai, M. Onishi, A. Kojima, and K. Fukunaga: Learning and Recognizing Behavioral Patterns Using Position and Posture of Human Body and Its Application to Detection of Irregular States, *Systems and Computers in Japan (SCJ 2005)*, Vol. 36, No. 13, pp. 45–56, (Feb. 2005).
6. M. Kourogi and T. Kurata: A method of personal positioning based on sensor data fusion of wearable camera and self-contained sensors, *In Proc. of IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2003)*, pp. 287–292, (Aug. 2003).
7. T. Maekawa, Y. Yanagisawa, Y. Kishino, K. Ishiguro, K. Kamei, Y. Sakurai, and T. Okadome: Object-based Activity Recognition with Heterogeneous Sensors on Wrist, *In Proc. of the 8th International Conference on Pervasive Computing (Pervasive 2010)*, pp. 246-264 (May. 2010).
8. Ulf Blanke, Robert Rehner and Bernt Schiele: South by South-East or sitting at the desk. Can orientation be a place?, *In Proc. of 15th IEEE International Symposium on Wearable Computers (ISWC 2011)*, pp. 43–46, (Jun. 2011).
9. T. Okuma, M. Kourogi, N. Sakata, and T. Kurata: A Pilot User Study on 3-D Museum Guide with Route Recommendation Using a Sustainable Positioning System, *In Proc. of the 13th International Conference on Control, Automation and Systems (ICCAS 2007)*, pp. 749–753, (Sept. 2007).
10. K. Willis, I. Poupyrev, T. Shiratori: MotionBeam: A Metaphor for Character Interaction with Handheld Projectors, *In Proc. of the 28th ACM Conference on Human Factors in Computing System (CHI 2010)*, pp. 1031–1040, (Apr. 2010).
11. T. Nakamura, T. Ogawa, K. Kiyokawa, H. Takemura: User Context Recognition for Use in Wearable Learning Systems Based on Congestion Level Estimation of the Inside of a Train Using a Carbon Dioxide Sensor, *ICEC Technical Committee on Multimedia and Virtual Environment (MVE 2007)*, Vol. 107, No. 554, pp. 49–54, (May. 2008).
12. R. Gross, L. Sweeney and F. Torre, S. Baker: Model-based face de-identification, *In Proc. of the 6th IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW 2006)*, pp. 161–168, (Jun. 2006).
13. J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma: Robust Face Recognition via Sparse Representation, *In Proc. of the 12th IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI 2009)*, vol. 31, Mo. 2, (Jan. 2009).
14. P. Viola and M. Jones: Rapid object detection using a boosted cascade of simple features, *In Proc. of the 15th Computer Vision and Pattern Recognition (CVPR 2001)*, pp. 511–518, (Dec. 2001).
15. OpenCV: `http://opencv.jp/.`
16. Wearable Toolkit: `http://wearable-toolkit.com/index.html`.
17. M. MIYAMAE, T. TERADA, Y. KISHINO, M. TSUKAMOTO, and S. NISHIO: An Event-DrivenWearable System for Supporting Motorbike Racing Teams, it In Proc. of the 8th IEEE International Symposium on Wearable Computers (ISWC 2004), pp. 70–76, (Oct. 2004).
18. M. MIYAMAE, T. TERADA, Y. KISHINO, M. TSUKAMOTO, and S. NISHIO: An Event-driven Navigation Platform for Wearable Computing Environments, it In Proc. of the 9th IEEE International Symposium on Wearable Computers (ISWC 2005), pp. 100–107, (Oct. 2005).
19. K. Tanaka, M. Kazuya, S. Nishio, S. Tanaka, K. Kinoshita, Y. Minami, T. Terada, M. Tsukamoto: IT-enabled donation boxes to promote donation, *In Proc. of the 7th International Conference on Advances in Computer Entertainment Technology (ACE 2009)*, pp. 400–403, (Oct. 2009).
20. videoInput: `http://muonics.net/school/spring05/videoInput/`.
21. com0com: `http://com0com.sourceforge.net/`.