Context–Aware Multi–Modal Notification for Wearable Computing

Nicky Kern and Bernt Schiele
Perceptual Computing and Computer Vision, ETH Zurich
{kern,schiele}@inf.ethz.ch

Abstract

We propose to use context information obtained from body–worn sensors to mediate notifications for a wearable computer. In particular we introduce a model which uses two axes, namely personal and social interruptability of the user in order to decide both whether or not to notify the user and to decide which notification modality to use. Rather than to model and recognize the complete context of the user we argue that personal and social interruptability can be derived directly from various sensors by the combination of tendencies. First experimental results show the feasability of the approach using acceleration, audio, and location sensors. The investigation of user interface issues are the immediate next steps in our work.

1 Background of the Author

Nicky Kern is a PhD student since October 2001 in the Perceptual Computing and Computer Vision Group (PCCV) at ETH Zurich. His work is supervised by Prof. Bernt Schiele.

His research focuses on using non–trivial context information obtained from body–worn sensors for context–aware applications (see [2]). In his current work he tries to use context information to mediate notifications from various sources to the user of the wearable computer.

2 Introduction

With the increasing number of wearable devices used by people in their everyday lives, there is an equally increasing number of applications that aim to grab the user's attention by various notifications. Be it arriving e–mails or telephone calls, upcoming meetings, changes in the stock market or navigation directions, the list of notifications on a wearable computer that can happen anywhere at any time in any situation is increasing. Clearly, there is a need to carefully handle and manage these increasing number of notifications in order to prevent wearable devices to become highly annoying. Importantly, management of notifications should take into account that the value of receiving a notification varies depending on the user's context. Also, since wearable computer and devices may offer a variety of notification modalities the notification modality can and should be chosen depending of the context and social situation of the user.

Any notification has two sides for the user: on the one hand it has a value and on the other hand it comes at the cost of interrupting the user. It is well known that interruptions can decrease work performance considerably. In everyday life, the user's primary task is often unrelated to the wearable computer, such that interruptions can be highly annoying and may be even dangerous.

We differentiate five important factors of the current user context: the *importance* of the event, that is being notified, the *activity* of the user, the *social activity* (if the user is interacting with others and if so, in which way), the *social situation* ('in a restaurant', etc.), and finally the *location*. It is important to point out that none of these factors is sufficient alone to make the best possible decision. The first factor of the five, the importance of the event, defines the value of the notification to the user. It is mostly unrelated to the user's context. Since its automatic determination is a research topic on its own, we assume it as given for the purpose of this paper.

There exist three principal ways to obtain the remaining four factors: they can be supplied by the user, be infered from usage patterns or acquired automatically using sensors. The first possibility is clearly undesirable since it does not reduce the load of the user. Since the second possibility can never supply us all the of the four factors, we propose to use multiple, body–worn sensors to acquire the necessary context information. Hudson et al. [1] have proved this to be possible.

In today's context-aware computing it is common to define or describe special situations or contexts which are directly related to the behavior of a 'context-aware' application, such as 'working in your office'. Such situations are however unsuitable for notification management. We therefore concentrate on lower level context information, such as 'user is having a conversation', which is more directly

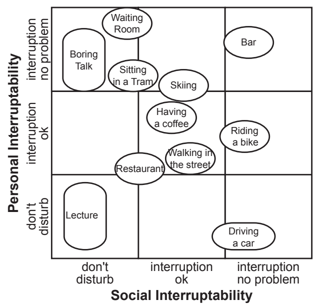

Figure 1. The Design Space of Notification

linked to notification management.

This paper is a summary of work that is going to appear at ISWC 2003 [2]. We first give a summary of the methods and experiments (sections 3, 4, and 5) and then address the goals of the workshop (section 6).

3 Design Space of Notification

When considering notifications in a mobile setting, there are three principal aspects to consider. Firstly the current user activity and social situation, which together determine whether to notify or not (section 3), secondly the most appropriate notification means (section 3.2), and finally the possibility of user intervention in exceptional situations (section 3.3).

3.1 Spanning the Design Space

Notifications are used to notify the user about an event that has happened, such as incoming phone calls or new e–mails. The event generally has a certain importance for the user. The notification itself has two sides for the user: on the one hand it has a value, because it conveys some important information about an event. On the other hand, it has a cost, because it interrupts the user in his current task. The value depends directly on the importance of the event to the user. As explained before in this paper, we focus on estimating the cost and assume the importance as given.

The cost depends directly on the degree of interruption. The degree of interruption depends on the interruptability of the user. This is highly depend on the current situation and activity of the user: e.g. while having a drink in a bar with some friends, a private call from the one missing friend would be appreciated, while the same call on the way home, while driving a car, would be much more interruptive. Therefore the *interruptability of the user* (referred to as *personal interruptability* in the following) is the key to evaluate the cost of a notification. It can be seen as a continuous variable that ranges from the extreme *Don't Disturb* over the intermediate range *Interruption OK* to the other end of the scale *Interruption No Problem*.

A notification does not necessarily reach the user only. An audio alarm can also be perceived by the environment — a potentially embarrassing situation, e.g. in a lecture. Thus we distinguish between the interruptability of the user (*personal interruptability*) and that of the environment (*social interruptability*). This allows us to choose a means of notification that interrupts the user only or interrupts both the user and the environment.

Our model consists thus of the two–dimensional space that is spanned by the personal interruptability and the social interruptability. Figure 1 shows the space with some example situations. The activity of 'Driving a car' requires much attention by the user, which has thus a low personal interruptability. It would however not disturb others, if he was notified, thus the social interruptability is high. For the situation 'Boring Talk' or 'Waiting Room' things are reversed: it would be highly unacceptable to notify the environment, e.g. using a loud ring, however an interruption of the user would probably be no problem if not appreciated.

3.2 Multi-Modality & Notification Intensity

Notifications are not binary, some allow to grab only part of the user's or the environment's attention. They can carry different amounts of information, they can be conveyed using different modalities, and some devices also allow to scale their interruptiveness (e.g. audio by changing the volume). The right notification modality should be chosen depending on the personal and social interruptability. We propose a scheme that allows to classify a notification's intensity and facilitates the matching from personal and social interruptability.

The intensity of a notification can be scaled from not notifying at all to trying explicitly to grab the entire user attention. We have to distinguish between the intensity for the user and the intensity for the environment. For example, a notification can be intense for the user, such as a flashing head mounted display (HMD), and completely imperceivable to the environment. We can thus classify notification modalities according to their intensity. Figure 2 shows the classification of two example sets of notification devices.

Since it becomes harder to ignore a notification with increasing intensity, intense notifications can only be used when the user or the environment have enough attention available. This corresponds directly to the interruptability we are using in the Design Space of Notifications. Hence we can discretize the Design Space of Notification and make a simple one–to–one mapping to choose the best notification modality depending on the personal and social interruptability.

The intensity of a notification also varies with the amount of information conveyed. The amount of information can be varied from a simple binary pulse 'something has happened' to very complex information, such as the complete text of an e–mail. Depending on the device this can be a near–continuous space in which any amount of in-

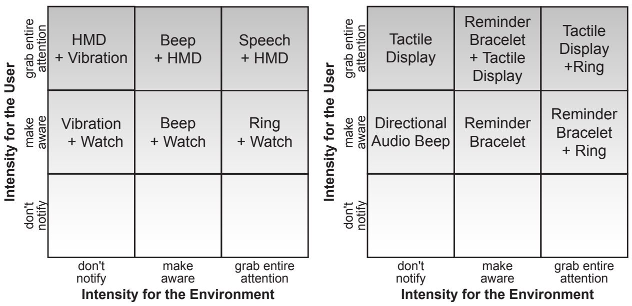

Figure 2. Example Configurations for 2 Sets of Devices (Left: HMD, Watch with Display, Audio, Vibration. Right: (Directional) Audio, Tactile Display, Reminder Bracelet.

formation can be conveyed or, for simpler devices, one of several distinct levels. Since the user needs more attention to deal with an information–rich notification, the intensity of a notification increases with the amount of information conveyed. In figure 2 the grey level indicates the amount of information that can be conveyed with a given notification.

Depending on the available devices this space has to be populated differently. Figure 2 shows two examples of two different sets of devices, that might be available to a user of a wearable computer. The set on the left consists of an HMD, a Watch with a display, a vibration device, and an audio device. The second set of a directional audio of a tactile display, a directional audio system, a reminder bracelet, and a normal audio device.

3.3 User Control

An important aspect is that the user must be able to control the way he is notified. The Design Space of Notifications is simple enough that users can directly choose their personal and social interruptability by hand, if need be.

As opposed to the direct configuration of the notification modality, this allows to select the best notification modality automatically, depending on the available devices. If the user left the HMD at home, he would not have to configure that he does not want notifications on the HMD. Instead the system can choose the best available notification modality automatically, depending on the available devices.

4 Estimating the Interruptability

In order to estimate the interruptability of user and environment, we propose to combine the inputs from all sensors and the prior that is given by the user. As pointed out in section 2 we do not model or recognize entire situations, but estimate the interruptability directly from the sensors, which give us a *tendency* for the interruptability.

When considering a single sensor only, we can already infer something about the user's and environment's interruptability. In the social situation 'Lecture', the user is probably personally little interruptible and quite certainly socially not interruptible. The sensor gives us a *tendency* which is the most probable interruptability. This tendency can be any kind of function that return the likelihood of interruptability on the entire Design Space of Notifications.

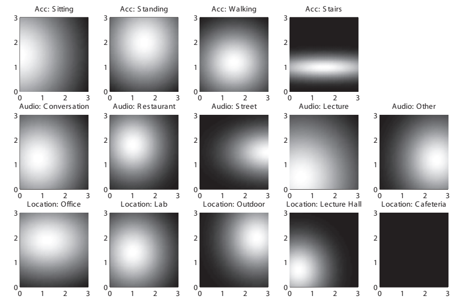

Figure 3. Tendencies for the sensors of section 5

Figure 3 shows the tendencies we chose for the experiments in section 5. Every tendency is represented by a 2–dimensional Gaussian. While activities such as 'Sitting' or 'Standing' can occur in nearly any situation, they cover a large area. However activities such as 'Stairs' or social situations such as 'Lecture' imply very specific interruptabilities.

The next step after defining tendencies for all sensors is combining them to obtain a final estimation. Since our 'sensors' are not physical sensors, but rather classification sub–systems, we can obtain for every 'sensor–reading' a likelihood. We use this likelihood to combine the tendencies: they are weighted according to their classification likelihood and summed. This is a preliminary step for the final estimation and called the *sensor estimate*.

The final part to take into account is the user's prior. As explained in section 3.3 the user can set his personal and social interruptability manually. This can be seen as an additional tendency. However, it has to be treated separately, because it has to have the possibility of overriding all other tendencies. The sensor estimate is weighted with a confidence measure in its correctness, e.g. using its variance or entropy. The weighted sensor estimation is combined with the user's prior to produce the final estimate of personal and social interruptability.

5 Experiments

To show the feasibility of our approach we have conducted an experiment that includes audio and acceleration context, as well as location context from wireless LAN access points, in order to estimates the social as well as the personal interruptability of a user from them.

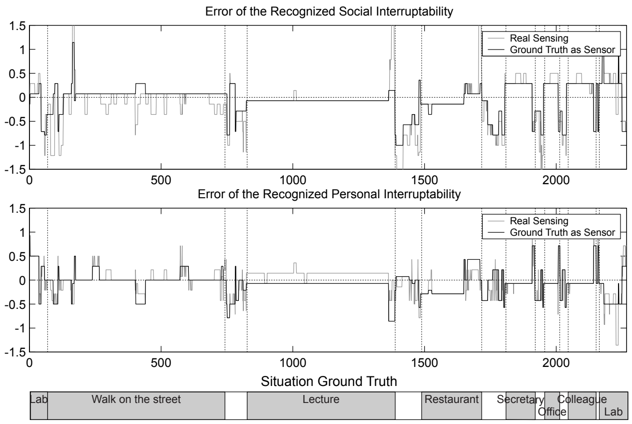

Figure 4. Error of Interruptability Estimation

Experimental Setup. We recorded a 37min stretch of synchronized audio and acceleration data. The data contains a walk on the street an visit to a lecture and to the computer science cafeteria as well as several discussions. Interruptabilities were manually annotated after the recording. Acceleration context was classified as *sitting*, *standing*, *walking*, *stairs* using a single sensor. Audio context was classified as *street*, *conversation*, *lecture*, *restaurant*, and *other*. Location was acquired from Wireless LAN access points and clustered into five different locations *Office*, *Outdoor*, *Lecture Hall*, *Lab*, and *Cafeteria*. Figure 3 shows the tendencies that were used to combine the contexts.

Results. Figure 4 shows the error of the interruptability estimation for social and personal interruptability. The interruptabilities are in the range $[0;3]$ as in figure 1. The axis at the bottom indicates the situation the user was in.

It is important to note that although the space is continuous, the selection of notification devices requires a discretization of the space. Using the discretization bins as in figure 2, we can tolerate errors up to 0.5 in either direction.

In order to verify that using tendencies is a sensible approach, we used the context ground truth as sensors and estimated the interruptabilities. The dark lines indicate these results. The error of the social interruptability estimation is below 0.5 and thus sufficiently precise for 88.5% of the time resp. 96.3% for the personal interruptability.

The grey lines indicate estimation results using real sensors. The error is 96.2% of the time below 0.5 for the personal interruptability and for 86% of the time for the social interruptability. Since the latter depends more on the audio context, its higher error rate introduces the higher error.

6 Workshop Goals

Devices. Our notion of a wearable computer is that of a collection of cooperative devices. The user might not always carry all his devices on him, but might select according to device capabilities, need or personal preference. Thus, we seek solutions that allow to adapt to a varying set of different devices.

Enabling Software Techniques and Architectures. As we have pointed out in the introduction, we have identified five factors of the user's context as important. The *importance* of the event that is being notified, the *activity* of the user, the *social activity*, and the *location*. All these five should be taken into account for the best possible decision. In this paper, we have concentrated on estimating the last four.

Design of Multi–Device Interfaces. We have introduced the Design Space of Notification that can be used for peripheral information display as well as notification. Both applications are highly interrelated and cannot be separated.

We have not addressed the issue on how the interaction after the notification will proceed. Currently we assume some high–bandwidth device, such as a PDA or a Head–Mounted display (HMD) available.

7 Conclusion and Outlook

Managing notifications is an important problem for future wearable computing devices and applications. They should be mediated automatically in order to avoid user annoyance. To this end we propose to use context acquired from multiple, body–worn sensors.

We have introduced a model that classifies notifications according to the *personal* and *social interruptability* of the user. It allows both to handle notifications in a systematic manner and to enable the automatic selection of the best notification modality.

Preliminary experiments have shown the feasibility of the approach experimentally. The context extraction seems to deliver sufficiently good results (see also [3]). The immediate next steps include the investigation of user interface issues and the usability of the social/personal interruptability model as interface to the user.

References

[1] S. Hudson, J. Fogarty, C. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. Lee, and J. Yang. Predicting human interruptibility with sensors: a wizard of oz feasibility study. In *Proceedings of the conference on Human factors in computing systems*, pages 257–264. ACM Press, 2003.

[2] N. Kern and B. Schiele. Context–aware notfication for wearable computing. In *Proc. 7th International Symposium on Wearable Computers (ISWC)*, October 2003.

[3] N. Kern, B. Schiele, and A. Schmidt. Multi–sensor activity context detection for wearable computing. In *Proc. European Symposium on Ambient Intelligence (EUSAI)*, Eindhoven, The Netherlands, November 2003.