WatchMe: communication and awareness between members of a closely-knit group

Natalia Marmasse, Chris Schmandt and David Spectre

MIT Media Laboratory
20 Ames Street, Cambridge MA02142, USA
{nmarmas, geek, spectre}@media.mit.edu

Abstract. *WatchMe* is a personal communicator with context awareness in a wristwatch form; it is meant to keep intimate friends and family always connected via awareness cues and text, voice instant message, or synchronous voice connectivity. Sensors worn with the watch track location (via GPS), acceleration, and speech activity; this is classified and conveyed to the other party, where it appears in iconic form on the watch face. When a remote person with whom this information is shared examines it, their face appears on the watch of the person being checked on. The working prototype was used as the focus of interviews to gauge the desirability of such a device.

WatchMe is a watch-based personal communicator that draws upon features of both mobile telephony and context-aware ubiquitous computing and integrates them in a user interface that is novel to both these domains. *WatchMe* extracts information from sensors to provide awareness and availability information to one's closest friends. It supports multiple modes of verbal communication (text messaging, voice messaging, and synchronous voice communication) enabling the recipients of the awareness information to choose the best communication modality. Photographs serve as emotional references to our loved ones, appearing on the watch when one of them is thinking of us.

1   Motivation

Everyone has a small group of people with whom they are emotionally close, a set of people who are very important in their lives. These are typically family members and/or intimate friends; people from our "inner circle" whom we call *insiders*. Nothing can replace the richness of face-to-face communication with these people; however, with our ever mobile and hectic lives, that is not always possible. Our aim is to use mobile communication ubiquitous computing to enable these people to keep in contact with each other. We would like to increase and facilitate communication, in a variety of modalities, among these small sets of intimate people. It is our hypothesis that people would want communication with this very select group of dear people

everywhere and all-the-time, as long as it were not too intrusive and they felt in control. We built a working prototype to demonstrate its feasibility and provide a focus for evaluation of and discourse about the technology. Our system has different layers of information that afford different degrees of communication.

awareness: Awareness is based on sending some basic information about ones activities. This information must require very low bandwidth since the system is always on, and hence constantly sending a trickle of data. We find that for awareness data to be meaningful at a glance, it must be abstracted; it requires more effort to interpret raw sensor data, so we don't display it. The person receiving our context data is not a stranger, but rather someone who knows us well and therefore can help interpret properly abstracted sensor data. The awareness data must be both collected and abstracted automatically; we simply do not believe that people will update it manually. This awareness data is the background information layer. A person is going about his way, sending out this awareness data to his intimates, having no idea if anyone is paying attention to it.

"thinking of you": This is the second layer of information, and the next layer up in terms of (tele)communication intimacy. The information being sent from one side causes changes to the display on the other side, i.e. person B is made aware that person A is thinking of him. At this stage there has not yet been any formal communication or exchange of verbal messages. This information transfer must require low bandwidth and have a low level of intrusiveness.

message exchange: After checking availability, or in response to "thinking of you", one party sends a message. There are three levels of messages.
- asynchronous text (e.g. text instant messaging)
- asynchronous voice (e.g. voice instant messaging)
- synchronous voice (e.g. full-duplex phone call)

These different modes of messages are increasingly intrusive. The system should enable a person to make an informed decision regarding the mutually preferable mode of communication. Escalation of the mode can occur during the flow of the communication. For example, if a person sees that another is thinking about them, they might respond by sending a message saying "want to talk?", or alternatively "I'm really busy!".

We find that such a system has four basic requirements. First, it should be always with you and always on. Second, the awareness data must be automatically gathered. Third, the system must be able to alert the user in subtle ways –the user needs to be aware of the awareness information if paying attention or not focused on some other task. Finally, it must be able to support communication modalities with multiple degrees of intimacy –i.e. different media.

After considering many alternatives, we selected a combination of a mobile phone and sensors built into a watch (Fig. 1). We strongly believe in the importance of a working prototype both as proof of concept, and to understand the technical difficulties and feasibility of the system. We have found the prototype to be invaluable for evaluation and to engage dialog about the different aspects of the

project, both amongst ourselves and with other colleagues or test subjects. We consider evaluation to be a multi-phase process: there is an evolution (of form and function) based on internal critique; we are influenced from our own and other people's investigation of user requirements for such technology [18, 14]; and evaluation continues through user studies and small focus groups.

In this paper we describe *WatchMe*, a mobile communication and awareness platform embodied in a watch. We describe the system hardware, functionality and user interface, including evolution of the design, and situate it in related work. We recount feedback received in a user interface evaluation and a pilot survey we conducted to assess peoples' acceptance of such a technology. Finally, we discuss privacy issues for such a device.
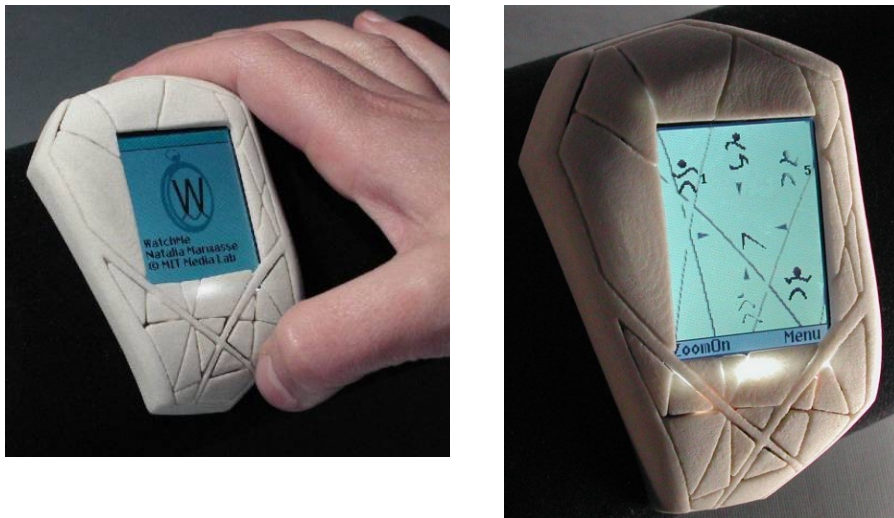

Fig. 1. *WatchMe* prototype displaying the main screen (right). Left image shows size of the current version.

1.1   Why a watch?

A watch is an artifact very assimilated into our lives. It is something most people wear, something we glance at numerous times a day. It is always accessible, always on, and in the periphery of our attention. Watches are very noticeable, but in a non-intrusive manner.

The device had to include mobile phone capabilities since one can hardly imagine a system for intimate telecommunication that doesn't include duplex synchronous voice. From a telephone network point of view text messaging, asynchronous voice and synchronous voice may be handled in very different ways. However from the user's point of view, they are all just different ways of reaching the same person, with different levels of intimacy.

Building such a system into a watch is a challenge, due to its physical size. A key requirement of the user interface is that it must convey a lot of information in a

relatively small amount of space, and in an aesthetically pleasing manner. An additional requirement was a device that could comfortably support switching between the modalities. A watch is in a location that is easily manipulated –albeit with one hand.

2   Hardware

The hardware comprises three components: the display and user input, the communication radio unit, and the sensing and classification unit. Our initial design rationale required that the user interface be easily accessible and frequently visible, which lead to a watch-based design. But to date appropriately sized hardware is not available, nor could we build such tiny phones. Although we see a rapid evolution of phones (display, processing power, size) such that a watch is a reasonable hardware target, we were forced to build our prototype with separate components. This is actually consistent with an alternative hardware architecture with several components, in different locations on or near the body, that communicate via a low power Personal Area Network, such as Bluetooth.

We would like to emphasize the three components of our prototype themselves, since the interconnections between them, although adequate for proof of concept, would have to be refined in a commercialized version.
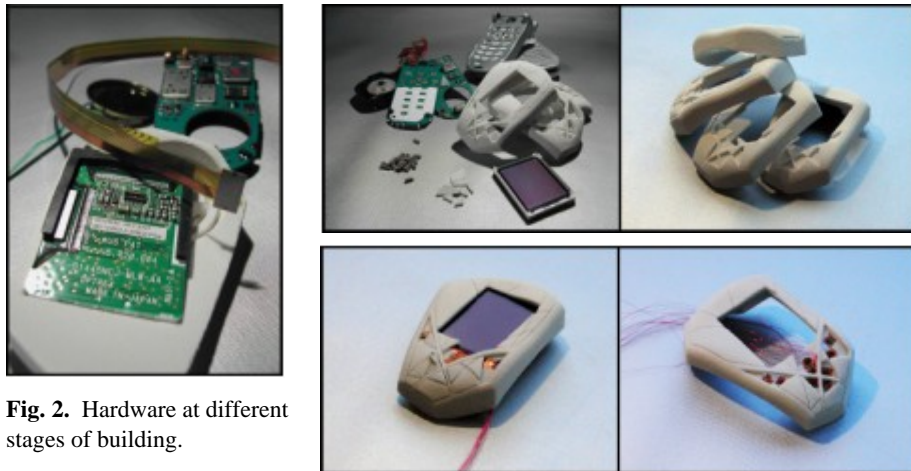

Fig. 2. Hardware at different stages of building.

display and user input: The display was removed from a Motorola iDEN mobile phone and encased in a shell built using a rapid prototyping 3D printer. This same shell includes the buttons for the user input, and is generally (together with the UI) what we refer to as "the watch". At this point the internals of the phone aren't in the watch. The display and buttons are tethered to the base of the phone, i.e. the communication component, via a flat flex cable and thin wires (Fig. 2). The watch shell also contains a speaker and microphone.

wireless communication: The radio component is the base portion of an iDEN phone, i.e. with the display part of the clamshell removed. It is connected to the watch component via a flat flex cable and wires. iDEN is a specialized mobile radio network technology that combines two-way radio, telephone, text messaging and data transmission in one network. It supports an end-to-end TCP/IP connection, the only platform that did so when we initiated this work. Other networks, such as GSM/GPRS, could also support our watch, with a different radio unit. The *WatchMe* system supports text messaging as well as voice messaging, using TCP/IP sockets. It also supports synchronous voice communication, using the ordinary mobile phone telephony functions. In this prototype the phone can be up to 35cms from the watch, limited by the length of the flex cable, so it could be strapped to the user's forearm.

sensing and classification: This component is made up of sensors, connected to or embedded in, an iPaq PDA. The iPaq reads the sensors, does data collection, and classifies the input. The current prototype includes three sensors: a Global Positioning Sensor to classify locations, an accelerometer to classify user activity, and a microphone for speech detection. The iPaq is clipped to the user's belt. The GPS unit can be embedded in the phone or connected to the PDA.

3   Functionality

The system can be divided into three different functional components: the watch, which comprises the user interface and display; the radio, through which the wireless communication is established; and the sensors and classification component, from which the personal context data is abstracted. There is also a server, which simply relays messages and context data from one user to another.

3.1   Watch User Interface

A watch is a personal device, but it is also very public. We often look at other people's watches to know the time when it would be socially awkward to look at our own. Watches are also often a fashion statement, meant to be looked at by others. Since it is at the seam of the personal and the public, the interface has tiers of different levels of information, with different levels of privacy.

The face of the watch is visible to all and conveys information accessible to all, i.e. time. People glance at their watch more often than they perceive. By embedding this high-level information in the watch's default mode, we can keep track of our loved-ones subconsciously and continually throughout our day. The top level, the default screen, also embodies other information meaningful only the owner. The owner of the watch chooses a unique icon and position around the watch face for each insider; although this is visible to others, they do not know the mapping from icons to names. Research has shown [18] that with text messaging clients, users interact recurrently with 5-7 people on a general basis. To play it safe, we chose to display icons for up to eight insiders. At this top level the colour of the icon indicates availability, fading to

the background colour in 3 steps: the most faded colour indicates that this insider does not have cellular coverage, the midway colour indicates that the person is in a conversation and hence probably less available. Speech is indicative of social engagement, and it has been found to be the most significant factor in predicting availability [17], therefore it was coded into the top level screen.
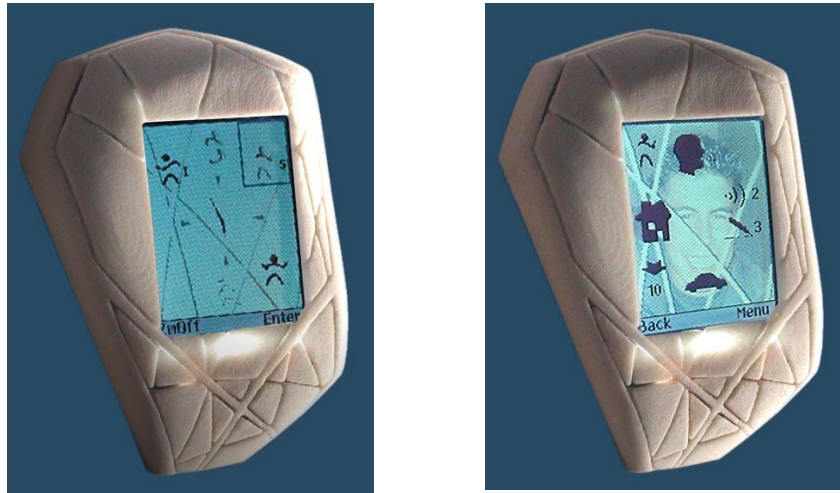

Fig. 3. Screen (left) showing cursor positioned on icon of insider. Pressing the Down navigational button will bring up the more detailed context information screen (right).

From a full-colour icon it is not possible to infer availability without going down a level in the interface and seeing more detail (Fig. 3). This is done by selecting the corresponding icon, via the Left/Right navigational buttons, and then pressing the Down button. On this screen a pre-selected image of the insider appears lightly underplayed in the background, as do the continuous lines of the design.

The more detailed information that can be viewed here (described clockwise from the top left) is the specific person's assigned icon, whether that person is engaged in a conversation, how many voice and text messages this person has left, and the person's mode of transport (walking, vehicle, biking, etc). Also displayed is his current location or next predicted one and expected time of arrival, or his last known location and time elapsed since departure. For example, in Figure 3, we see that Joe left home 10 minutes ago, that he is driving and in a conversation, and that he has sent 2 voice messages and 3 text messages; the top level shows that he has left 5 messages total. Although it is necessary to navigate to this screen for the detailed information, the top level provides an overview of all insiders, displaying salient information regarding their availability, and the number of new messages they have sent.

Since Joe is driving and also talking, this is probably not a good time to phone him. For an insider, this little information can go a long way. With a combination of prior knowledge and a form of telepresence provided by the watch, it is possible to quickly form a meaningful interpretation. For example, knowing Joe and judging by the time and that he is driving and talking, it is possible to presume that he has already picked up his buddy and is heading to the gym. If "gym" is a location Joe has revealed, once

the system has enough information to predict he is heading there, the icons will change to reflect that (gym icon, direction arrow, and ETA).

The watch supports text messaging, voice messaging, and phone calls. The content of the text and voice messages, as well as the ability to compose messages or place a phone call to the insider, is accessed through yet a deeper layer.

A fundamental part of communication is its reciprocal characteristic. When an insider lingers viewing another's detailed information (in this case, that she is biking to work and expected to arrive in 18 minutes), her image appears on the reciprocal wristwatch (Fig. 4). In this way one can have a notion of when a specific insider is thinking of the other, and this information may subsequently stimulate an urge to contact that person. This conviction is supported by [18] where a significant fraction of the communication happened immediately after a party appeared online.

Knowing that someone is thinking of you creates opportunity for communication, but not obligation. When the picture appears on the "viewed" insider's watch, one of the following could occur:

- The picture popping up may go unnoticed, especially since it disappears after a couple of minutes, so the "viewing" insider is not interfering with the "viewed" insider in any way.
- The "viewed" insider notices the picture but decides not to reply or divert attention from his current action.
- The "viewed" insider notices the picture and responds by querying the availability of the other user, which causes his or her picture to appear on the other's watch, similar to an exchange of glances without words.
- The "viewed" insider decides to phone the "viewer" or engage in another form of verbal communication, i.e. text or voice messaging.

Fig. 4. When an insider thinks about another and views her detailed context data (left), the "viewer's" photograph will appear on the "viewed" insiders watch (right).

There are a number of alerting modes on the watch. For example, when your picture appears on my watch, indicating that you are thinking about me, the backlight

turns on to draw a little attention. The watch can also vibrate or emit sounds and these could be used as well if the wearer wants the watch to be more intrusive. These same features can also be used for non-verbal communication. When a picture appears, the user can send back a photograph from a stored repository, or alternatively manipulate the other individual's watch backlight or vibration actuator enabling them to develop their own non-verbal codes.

The user interface design has been a continual process. It is important that there be harmony between the graphics on the screen and the physical form of the watch itself. Figure 5 shows some previous designs of the interface and drawings of other forms considered for the watch.
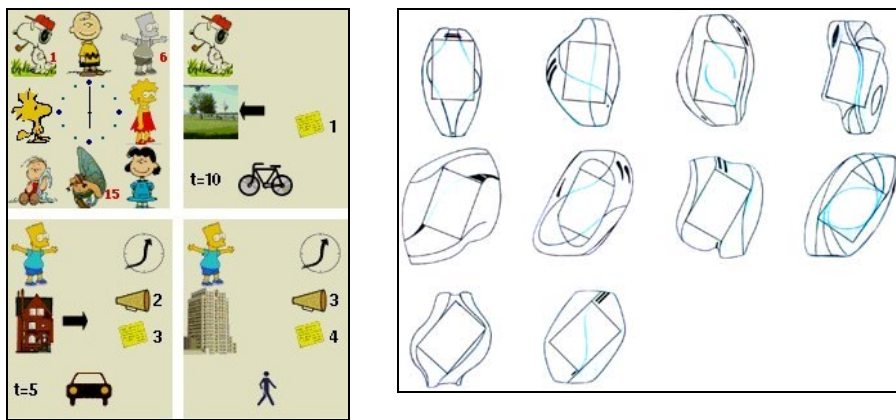

Fig. 5. Left: the first UI design. Right: hand sketches exploring preliminary variations of shape and screen rotation.

3.2 Radio

The radio component (an iDEN phone without the screen or buttons) is connected to the display and buttons in the watch. This unit performs the processing required for the user interface, manages the socket connection to the server (which relays the messages between users), and performs the telephony functions required for the synchronous voice communication.

All mobile phones have microphones, many already have embedded GPS chips (at least in the U.S. due to the FCC E-911 wireless location mandate), and soon some will have embedded accelerometers –these can also be connected via the phone's serial port. So although this unit could encompass all the required sensors, the limiting factor is its computing power. Therefore the classification is performed on an iPaq, and the classifier outcome is communicated to the phone unit.

3.3 Sensors and Classification

Cues from the physical world often help us infer whether a person is interruptible or not. An office with a closed door, for example, may indicate that the person is not

around, or does not want to be disturbed. However from prior knowledge we may be aware that this particular person is easily distracted from outside noise and therefore keeps the door shut, but that it is perfectly acceptable to simply knock. If a door is ajar and voices can be heard, then perhaps the person is unavailable –that could depend on the nature of the relationship and the urgency of the topic. Throughout our lives we have acquired a whole protocol of what is appropriate in different (co-located) social contexts. How do we do this at a distance? What is the subset of cues necessary to convey to people (who know us well) that will help them infer our availability?

Locations are classified based on latitude/longitude, founded on an extension to the software from our comMotion system [20]. The original version detected frequented indoor locations, based on loss of the GPS signal. We have enhanced this model to also detect locations where the receiver is stationary with signal. When the system identifies a previously unnamed frequented location, it prompts the user to label it. In this way the system learns from and adapts to the user over time, only prompting him when an unknown location is encountered. The string associated to the labeled locations is what is reported to the other phones. A basic set of strings is associated with default icons, such as "home" and "work". A location will only be sent if it is named, and if the recipient hasn't associated an icon with that name, a text string appears instead. We also enhanced the comMotion model which analyses patterns of mobility to determine routes, positions along those routes, and an estimated time to arrival; a preliminary version of the algorithm was described in [21]. This is used to indicate, for example, that the user left home 10 mins ago, or will arrive at the office in 15 minutes.

GPS data over time allows velocity to be computed with enough resolution to differentiate between walking and driving (as long as not in urban gridlock). Although it is difficult to detect the difference between highway driving and riding a train, for example, the route classifier differentiates these two travel paths and the user has the ability to label them. For higher resolution classification, such as differentiating between walking, running, and bicycling, we rely on two orthogonal 2-axis accelerometers giving 3 axes of acceleration [23]; it is based on hardware developed jointly and a classifier similar to [3] which analyses the mean, energy, frequency-domain entropy and correlation between two different acceleration axes. With 5 sensors it is possible to correctly classify 20 activities such as walking, running, brushing teeth, folding laundry, and climbing stairs; *WatchMe* uses fewer degrees of classification.

The third sensor used is a microphone. Audio data, from the PDA's microphone, is collected and examined in near real-time to detect whether it is speech. The analysis involves taking 10 seconds of audio, looking at the pattern of the voiced segments in the pitch track, and determining whether it corresponds to speech. This is a binary speech discriminator, it is not necessary to know whether the speech is generated by the user himself or someone he is talking to; as he is probably in a conversation in either case. Likewise, we do not try to distinguish whether the conversation is over the phone or with someone physically present, though this could easily be determined. None of the audio is stored, nor do we try to perform any speech recognition.

Others have shown the value of sensors in identifying a person's context [7, 15], especially the determination of speech as a significant factor [17].

4 Privacy

In any awareness system some of the information that is revealed is sensitive to some of the participants at least part of the time. In the course of developing *WatchMe* we encountered a number of privacy issues.

sensitive information: *WatchMe* reveals a lot of information about a user, but only the locations that he has chosen to name; raw geographic coordinates are never revealed. A user might see that another is at the bookstore, but where the particular bookstore is physically located is not displayed. Additionally, *WatchMe* has been designed from the beginning to be a system used by people who are intimate friends. Since they already share much personal information, using technology to do so is less intrusive. People whom we are really close to know much more sensitive information about us than, for example, how long ago we left our house.

photographs: Photographs are very personal and a watch face is semi-public. People may be more sensitive in other cultures, but in ours we often display pictures of family, especially children, in offices and homes. We often carry them in wallets or purses, both to look at ourselves and to show to others. We now have them on phones as well, so displaying pictures of our loved ones on a watch is not that different. The detailed context information would not be readily understood by someone looking at our watch from a distance. It is also invoked only by specific user action.

reciprocity: *WatchMe* enforces reciprocity of data. A user cannot receive context data from another unless he is also sending his. There is also reciprocity of interaction: when user A views B's context data, A's photograph appears on B's watch. So a person can't "spy" on another without them knowing they are doing so, regardless of whether it carries a positive or negative connotation.

peer-to-peer vs. server: The current implementation depends on a server to relay the messages between the users. Now that there is better support of server sockets on the phones, the architecture could be modified to be peer-to-peer, over a secure socket, adding another layer of security. Even in this version, no data is stored on the server.

plausible deniability: The user has control over the locations he decides to share with his insiders, and at any given time he can manually make it seem that his watch is "out of service" (out of cellular range), or that he is in a conversation. We have thought about randomly invoking the "out of service" mode to provide the users with plausible deniability and prevent them from having to explain why suddenly they were disconnected. In this way it can be attributed to a supposed bug in the system, when in fact it is a privacy feature. The user's location is only transmitted to others when he is somewhere he has previously chosen to name, however the hardware that he is wearing is keeping a history of where he has been, to detect these patterns and perform calculations of ETA. In addition to giving the user the option of not sharing the location, he should also have the option of not logging it at all or the ability to delete certain sections from it. No acceleration data or audio is saved.

5   Related Work

A good deal of research has addressed how the awareness of presence, availability and location can improve coordination and communication. Much of it has focused on how to improve collaboration between work teams. Several systems require cameras and microphones set up in the workspace, as well as broadband connections, to support transmission of video and/or audio. Other systems require either infrared or radio frequency sensors, or heavy data processing. Recently there has been a focus on more lightweight systems for mobile devices –lightweight installation as well as easy to use. We will describe only a subset of all of these systems.

awareness through video and audio: The Montage [30] system provided lightweight audio and video "glances" to support a sense of cohesion and proximity between distributed collaborators. It used a hallway metaphor where one can simply glance into someone's office to see if it is a good time to interact. A similar metaphor was used in Cruiser [28, 11] which enabled a user to take a cruise around each office. The purpose of the system was to generate unplanned social interactions. In Portholes [8] non co-located workers were periodically presented with updated digitized images of the activities occurring in public areas and offices. Some systems have focused on awareness solely through audio. Thunderwire [1] was an audio-only shared space for a distributed group. It was essentially a continuously open conference call in which anything said by anyone could be heard by all. ListenIN [32] uses audio to provide awareness of domestic environments to a remote user. In order to add a layer of privacy, the audio is classified and a representative audio icon is presented instead of the raw data; if the audio is classified as speech it is garbled to reduce intelligibility.

location awareness: Groupware calendars have been useful tools to locate and track colleagues. Ambush [24] looked at calendar data to infer location and availability. It used a Bayesian model to predict the likelihood that a user would actually attend an event entered in his calendar. Calendars and Bayesian models have also been used to predict a user's state of attention [16]. Location-aware systems have used infrared or radio frequency sensors to keep track of electronic badges worn by people [33], or GPS [20]. The Work Rhythms project [4] looks at location of computer activity to create a user's temporal patterns. Awareness of these patterns helps co-workers plan work activities and communication. When a user is "away", the system can predict when he will be back.

context and mobile telephony: The so-called context-awareness of computer systems falls very short of what humans can assess. As Erickson [10] puts it: the ability to recognize the context and determine the appropriate action requires considerable intelligence. Several systems keep the human "in the loop" by enabling the potential recipient to select a profile appropriate for the context. In the Live Addressbook [22] users manually updated their availability status and the location where they could be reached. This information was displayed to anyone trying to contact them. Although the updates were manual, the system prompted the user when he appeared to be somewhere other than the location stated. Quiet Calls [26] enabled users to send callers pre-recorded audio snippets, hence attending a call quietly. The

user could listen to what the caller was saying and send a sequence of standard answers. Another system that shares the burden of the decision between caller and callee is Context-Call [29]. As with most profile options, the user must remember to update the stated context.

lightweight text communication: ICQ started as a lightweight text message web application in 1996. It has since grown into a multimedia communication tool with over 180 million usernames, and 30 million users accessing per month [2]. A user's availability is automatically set based on computer activity, however it can manually be overridden. Babble [9] aimed to support communication and collaboration among large groups of people. It presented a graphical representation of user's availability, based on their computer interaction. Nardi et. al. [25] studied the extensive use and affordances of instant messaging in the workplace. Desktop tools for managing communication, coordination and awareness become irrelevant when a user is not near their computer. Awarenex [31] extends instant messaging and awareness information to handheld devices. It has the concept of a "peek", an icon that appears in the buddy list indicating a communication request. Hubbub [18] is a mobile instant messenger that supports different sound IDs; the location data is updated manually.

non-verbal communication systems: There are also some systems that have looked at ways to enhance interpersonal communication by adding physical feedback via actuators. ComTouch [6] augments remote voice communication with touch. It translates in real-time the hand pressure of one user into vibrational intensity on the device of the remote user. The Kiss Communicator [5] enabled couples to send each other kisses. One person would blow a kiss into one side of the device and the remote piece would start to blink. The other person could respond by squeezing the communicator causing the lights to blink on the side of the original sender. The Heart2Heart [13] wearable vests conveyed wireless "hugs" by simulating the pressure, warmth and sender's heart-beat as would be felt in a real embrace. Paulos [27] suggests a system with sensors (accelerometer, force sensing resistors, temperature, microphone for ambient audio) and actuators (Peltiers, bright LEDs, vibrator, "muscle wire", speaker for low level ambient audio) to enhance non-verbal telepresence. This system will use Intel's Motes and will include a watch interface.

watches: Whisper [12] is a prototype wrist-worn handset used by sticking the index fingertip into the ear canal. The receiver signal is conveyed from the wrist-mounted actuator (electric to vibration converter) to the ear canal via the hand and finger by bone conduction. The user's voice is captured by a microphone mounted on the inside of the wrist. Commercial handsets built into wristwatches are also starting to appear, such as NTT DoCoMo's wrist phone or RightSpot [19].

6   Evaluation

We conducted both a pilot survey to assess peoples' acceptance of such a technology and a user study of the implemented user interface; we discuss those here. In this section we also discuss the next steps of the project.

6.1   Survey

The pilot survey, besides helping us understand which features are essential, has helped us put together a more comprehensive survey that is being conducted on a much larger scale. It was carried out on a group of 26 people spanning the ages from teens to sixty five, from four different countries (USA, Mexico, Israel and Sweden). The subjects were recruited by email, by people who worked on the project, and asked to answer an email questionnaire. The respondents were encouraged to forward the questionnaire to their friends. The vast majority of the subjects did not know about the project, but they were family or friends of friends of the researchers. The survey included two different scenarios and questions about them.

communication modalities and awareness: The first scenario asked the person to imagine s/he had a device, such as a keychain or mobile phone, which would enable their friends and family to know their whereabouts. The location information would be automatically available without any effort by either party, it would be reciprocal preventing one from "spying" on another, and a person would always have the option of switching the device off. It was pointed out that such a device would, for example, "enable a working mom to know that her husband had already left the office, that her son was still at guitar practice" (probably waiting to be picked up by dad), and that her daughter was already at home".

In this population, when face-to-face communication with family and friends is not possible, the most common alternatives are communication by phone or email, followed by text messaging (IM, SMS). The large majority would be willing to share information on their whereabouts only with immediate family, that is, spouse and children. A few would also share with close friends and siblings. Not surprisingly, some teens seemed much less enthusiastic about giving this information to their family, although an opportunity whereby the parents would be aware of inopportune moments to call was valued. People indicated that they would be willing to disclose locations such as: home, work, school, gym, supermarket, etc., but few would keep the device turned on all of the time.

feature set: New features people want included are: the ability to know who was watching you; the ability to talk to the person observing you; a "busy scale" which could either be set manually or "smartly" by the system; the ability to provide a false location if necessary; the option to leave messages; a "general" vs. "detailed" mode indicating for example "shopping" instead of the name of a particular store; the option to request a person to turn their device on; and preventing children from turning their devices off or overriding the system with a false location.

People definitely did not want the system to include: hidden cameras; the option for people to track you without your knowledge; the possibility of hearing everything said; the option to permanently store the information on a person's movements; and for unauthorized people to get a hold of this information. People were willing to give some location information to a few chosen people they trust, but were very concerned of being monitored without their consent and knowledge. Almost everyone said they would take into consideration a person's location before communicating with them, and would want this courtesy to be reciprocal. We asked what other information,

besides location, people would be willing to reveal. The responses received were very bimodal. Many people seem reluctant to provide more of their specific context information and prefer a more abstract "busy" or "do not disturb" label, whereas others want family trying to contact them to know that they are driving, or in a meeting, or on vacation, etc.

"thinking of you": The second scenario asked people to imagine a device that displayed a picture of whoever happened to be thinking about them. We wanted to know who people would be willing to share their thoughts with, so to speak, and how they would respond when the device displayed a picture of someone thinking about them. About 2/3 would share this experience with a combination of immediate family, close friends and siblings. One person said it would be nice to be able to let friends and family know that he was thinking of them without having to take time out to call or write a message, and that he could list at least 30 people he'd like to regularly let know he was thinking about them. About 1/3 found this idea "creepy" and did not like it. Of the group who liked the concept of the device, they would react to receiving a picture by: phoning the person if they were not too busy; have a "warm feeling", send them back a picture and maybe phone depending on who they were; would just be happy but not do anything about it; would respond only to spouse; or would email or call them to get together.

6.2   User Interface Evaluation

We conducted a small evaluation of our watch prototype, focusing on usability, choice of communication modes, and the appeal of such a watch. The 15 subjects (8 female, 7 male) were aged 25 to 48, including students and administrative staff and outsiders. The one-on-one sessions lasted from 20 minutes to 1.5 hours. First, we explained and demonstrated the user interface. Next subjects were given as much time as they wanted to explore the interface display and buttons; no subject spent more than two minutes doing so. Each subject was asked to perform 3 specific communication tasks using the device. The device logged the whole interaction and the subjects were observed while performing the tasks by one of the authors. At the end of the third task, each subject filled out a questionnaire. After completion of the questionnaire most of the subjects felt compelled to talk about the system in general and the prototype in particular, get more detail, and offer comments. Some of these unforeseen conversations over the prototype lasted close to an hour.

The first task was to send a text message to a specific person, the second task was to send a voice instant message to someone else, and the third task was to communicate in any modality to a third person. The first two tasks were directed at the usability of the watch, while in the third we wanted to see the utility of the context information of the remote person, and whether having that information affected the communication mode chosen.

usability: Subjects were asked on a 1-7 scale (1-very hard, 7-very easy) how easy the system was to use, and how well they thought they had performed. The mean and standard deviation for ease of use were $\mu = 5.67$ and $\sigma = 0.9$. For the self-reported

performance $\mu = 5.6$ and $\sigma = 0.91$, although the observer considered that all had managed to perform the task and everyone in 6-7 minutes total. Almost all the complaints related to the button interface, rather than system features or functionality. People found the buttons too small, making it hard to navigate. Some found it hard to distinguish the buttons and remember what they did, though this could be due to the novelty of the device. The robustness of the buttons was an issue, requiring us to re-glue them often. Some people liked that it wasn't clear which buttons were functional, making the watch look "more like jewelry, less nerdy". One way to reveal the buttons to the user only is to give them a slightly different texture. Clearly we will have to rethink and redesign the button interface.

A few subjects disliked the "texting". Text messges are composed by choosing characters from a soft keyboard, via the navigation buttons. Each chosen character is appended to the message string. Some users added more than one space character since they had no visual feedback that it had been appended. Once composed, the message is sent by pressing a different button. These two buttons were intentionally placed next to each other to facilitate quick texting with one thumb. Several users confused the buttons, sending incomplete messages. Although most didn't bother to send another message with the remainder of what they had intended to write, this could obviously be done. Perhaps only few mentioned these issues because texting on a small device is known to be problematic and hence their expectations were low.

choice of communication mode: In the third task, the person they were to communicate with had left them 2 text messages and 1 voice message; the context data indicated that she was driving and expected to be home in 35 minutes. 60% chose to give her a call with explanations such as: "she is driving so text is not a good option but she seems available"; "I called because her voice message said *give me a call*"; "it seemed urgent and this was the quickest way to reach her". Three people left a voice message and explained that the recipient was driving and therefore a phone call was not recommended, and three left a text message since it was the easiest for them. Seven said they considered the recipient's convenience, four considered only their own, one person considered both, and three considered neither.

The voice message the subjects listened to indeed said "give me a call when you get a chance", however this was said in a casual tone. Since the messages are from a fictitious person, and not from an insider as the system is envisioned to be used, the subjects' interpretation of the context varied. Those who thought it was urgent to get in touch with her did not believe convenience to be a relevant factor. One person misinterpreted the context data –he thought she had been home for the last 35 minutes, and not that her ETA was 35 minutes– he afterwards said that in that case he would have just waited until he saw that she had arrived home and only then phoned.

We also asked about general preferences of communication channels. Text messaging was the least preferred for sending but, significantly, what people said they preferred for receiving. Composing a text message on a small device with few buttons can indeed be tedious. The asynchronous text mode for reception is generally preferred since it can be accessed at any time and there are no privacy concerns with others listening in. It is also faster to read than to sequentially listen to audio.

appeal: Subjects were asked how much they liked the system (1-not at all, 7-very much), what they specifically liked and disliked about it, and who they would share this type of information with. People seemed to really like the system ($\mu = 6.07$, $\sigma = 0.96$); 2/3 would share this information with their spouse or boy/girl-friend, 7 would share with other family members such as siblings or parents, and 9 would share with some close friends. As to whether they would use such a system, 2/3 said "yes", 3 were undecided, and 2 said they probably would. However 6 said they would not wear it on their wrist (even assuming it was much smaller), and 2 were undecided.

As noted before, the predominant thing said against the system was the button interface. Many liked the icons and especially the information they convey: "this is the perfect device for me, often I call just to get the information that I can just see here". Someone noted that he would like the context data to feel more in touch with his girlfriend and other friends who are all on the other side of the Atlantic.

Comments regarding the different communication modalities were very positive: "I really like the features in this watch. In general I hate all-in-one devices, but this one is great. It groups together things that make sense, they all have to do with communication, and in a simple way"; "it let's me communicate more politely"; "I like the blurring of the boundaries between message types". Overall, subjects enjoyed the trial, found the technology stimulating, and wanted to talk about it at length afterwards. We find this very encouraging.

One surprising result was that seven of our subjects no longer wear watches. For some this is due to the physical constraints (heavy, make you sweaty, etc.), while many noted that the time is readily available, e.g. on their mobile phones, computers, or clocks in the environment. Clearly people who don't wear watches are less inclined to a technology that you wear on your wrist, but if the phone and the watch become the same gadget, this new trend may be reversed. In any case, a surprising number of people liked the technology; those who don't want it on their wrist would like to have a device you could clip to the belt or put in a pocket, or simply on a conventional mobile phone. Not having the device located in the periphery of visual attention would require rethinking the design of the interaction, perhaps relying more on auditory or tactile cues.

6.3 Future Steps

Except for completing the sensor integration, we have a fully functional prototype, in the shape of a wristwatch, built using a real phone. The watch is about 1.5 times the size we would eventually like it to be; new generation phones with their smaller screens will help reduce the size. We have identified a few problems with the current user interface; these will be addressed in the next version. While we don't claim *WatchMe* is suitable for everyone, a significant number of people who used it were very positive. More evaluation would be needed before it were made into a product.

Our previous work successfully evaluated both the GPS, location and route finders, and accelerometer-based classifiers. We do not yet have quantitative data as to the performance of the three classes of sensors (GPS, accelerometer, microphone) operating jointly, but since they are mutually independent we don't anticipate difficulties with fusion of these sensors. Nonetheless we will certainly evaluate the

classification component in the field. More importantly, we would like to evaluate how friends or couples would actually use *WatchMe* in real life. This requires robust enough engineering so that they can be taken out of the lab for periods of several months. We're especially concerned with issues of trust and confidence in security of the data between users who are intimate friends.

Our work is predicated in the belief of the importance of having a working prototype to properly evaluate people's response to the underlying concepts. For example, a person who in the survey had expressed some reservations about such a technology, was very enthusiast when she used the prototype in the user study. Each iteration of *WatchMe* has required new hardware and some engineering help from Motorola. We would like to express our sincere gratitude for their extensive support during the course of this project.

References

1. Ackerman, M., Hindus, D., Mainwaring, S., and Starr, B. (1997) Hanging on the 'Wire: A Field Study of an Audio-Only Media Space. *Transactions on Computer-Human Interaction* (TOCHI), vol. 4, no.1, pp.39-66.
2. AIM, Oct 2002. http://news.com.com/2100-1023-963699.html
3. Bao, L. (2003) Physical Activity Recognition from Acceleration Data under Semi-Naturalistic Conditions. M.Eng. Thesis Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Sept.'03.
4. Begole, J., Tang, J., Smith, R., and Yankelovich, N. (2002) Work Rhythms: Analyzing Visualizations of Awareness Histories of Distributed Groups. *Proceedings of the CSCW 2002 Conference on Computer Supported Cooperative Work*, pp. 334-343.
5. Buchenau, M. and, Fulton, J. (2000) Experience Prototyping. *DIS '00 Symposium on Designing Interactive Systems*, pp. 424-433.
6. Chang, A., O'Modhrain, S., Jacob, R., Gunther, E. and Ishii, H. (2002) ComTouch: Design of a Vibrotactile Communication Device. *DIS '02 Symposium on Designing Interactive Systems,* pp. 312-320.
7. Clarkson, B., Mase, K., and Pentland, A. (2000) Recognizing user context via wearable sensors. *Proceedings Fourth International Symposium on Wearable Computers*, pp. 69–76.
8. Dourish, P. and Bly, S. (1992) Portholes: Supporting Awareness in a Distributed Work Group. *Proceedings of the CHI '92 Conference on Human Factors in Computing Systems*, pp.541-547.
9. Erickson, T., Smith, D.N., Kellogg, W.A., Laff, M., Richards, J.T. and Bradner, E. (1999) Socially Translucent Systems: Social Proxies, Persistent Conversation and the Design of "Babble". *Proceedings of the CHI '99 Conference on Human Factors in Computing Systems,* pp. 72-79.
10. Erickson, T. (2002) Ask not for Whom the Cell Phone Tolls: Some Problems with the Notion of Context-Aware Computing. *Communications of the ACM*. Vol. 45, No. 2, pp.102-104.
11. Fish, R., Kraut, R., Root, R., and Rice, R. (1992) Evaluating Video as a Technology for Informal Communication. *Proceedings of the CHI '92 Conference on Human Factors in Computing Systems,* pp.37-48.
12. Fukumoto, M., Tonomura, Y. (1999) Whisper: A Wristwatch Style Wearable Handset. *Proceedings of the CHI '99 Conference on Human Factors in Computing Systems*, pp.112-119.
13. Grimmer, N., (2001) Heart2Heart. *Winner of Intel Student Design Competition 2001.* http://www.baychi.org/calendar/20001508/#1

14. Grinter, R. E. and M. Eldridge. (2001) 'y do tngrs luv 2 txt msg?', in W. Prinz, M. Jarke, Y. Rogers, K. Schmidt and V. Wulf (eds.): *Proceedings of the Seventh European Conference on Computer- Supported Cooperative Work ECSCW '01*, pp. 219-238.
15. Hinckley K., Pierce, J., Sinclair, M., and Horvitz, E. (2000) Sensing techniques for mobile interaction. *ACM User Interface Software and Technology*, CHI Letters 2: 91–100.
16. Horvitz, E., Jacobs, A. and Hovel, D. (1999) Attention-sensitive alerting. P*roceedings of UAI '99 Conference on Uncertainty in Artificial Intelligence,* pp.305-313.
17. Hudson, S., Fogarty, J., Atkeson, C., Forlizzi, J., Kiesler, S., Lee, J. and Yang, J. (2003) Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study, *Proceedings of CHI '03 Conference on Human Factors in Computing Systems,* pp. 257-264.
18. Isaacs, E., Walendowski, A., Ranganthan, D. (2002) Hubbub: A sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. *Proceedings of the CHI '02 Conference on Human Factors in Computing Systems,* pp. 179-186.
19. Krumm, J., Cermak, G., and Horvitz, E. (2003) RightSPOT: A Novel Sense of Location for a Smart Personal Object. *Proceedings of Ubicomp 2003: Ubiquitous Computing,* pp. 36-43.
20. Marmasse, N., and Schmandt, C. (2000) Location-aware information delivery with comMotion*. Proceedings of HUC2000 International Symposium on Handheld and Ubiquitous Computing*, pp. 157-171.
21. Marmasse, N., and Schmandt, C. (2002) A User-Centered Location Model. Personal and Ubiquitous Computing, vol 6, no. 5-6, pp. 318-321.
22. Milewski, A. and Smith T. (2000) Providing Presence Cues to Telephone Users. *Proceedings of the CSCW 2000 Conference on Computer Supported Cooperative Work*, pp. 89-96.
23. MunguiaTapia, E. and Marmasse, N. (2003) Wireless Sensors for Real-time Activity Recognition, http://web.media.mit.edu/~emunguia/portfolio/html/wsensors.htm
24. Mynatt, E. and Tullio, J. (2001) Inferring calendar event attendance. *Proceedings of the IUI 2001 Conference on Intelligent User Interfaces*, pp. 121-128.
25. Nardi, B., Whittaker, S, and Bradner, E. (2000) Interaction and Outeraction: Instant Messaging in Action. *Proceedings of the CSCW 2000 Conference on Computer Supported Cooperative Work*, pp. 79-88.
26. Nelson, L., Bly, S., and Sokoler, T. (2001) Quiet Calls: Talking Silently on Mobile Phones. *Proceedings of SIGCHI conference on Human Factors in Computing Systems*, pp. 174-181.
27. Paulos E. (2003) Connexus: A Communal Interface. http://www.intelresearch.net/Publications/Berkeley/070220031047_135.pdf
28. Root, R. (1988) Design of a Multi-Media Vehicle for Social Browsing. *Proceedings of the CSCW '88 Conference on Computer Supported Cooperative Work*, pp.25-38.
29. Schmidt, A., Takaluoma, A., and Mäntyjärvi, J. (2000) Context-Aware Telephony Over WAP. *Personal and Ubiquitous Computing*, vol. 4, no. 4, pp. 225-229.
30. Tang, J. and Rua, M. (1994) Montage: Providing Teleproximity for Distributed Groups. *Proceedings of CHI '94 Conference on Human Factors in Computing Systems,* pp. 37-43.
31. Tang, J., Yankelovich, N., Begole, J., VanKleek, M., Li, F., and Bhalodia, J. (2001) ConNexus to Awarenex: Extending awareness to mobile users. *Proceedings of the CHI '01 Conference on Human Factors in Computing Systems,* pp.221-228.
32. Vallejo G. (2003) ListenIN: Ambient Auditory Awareness at Remote Places, M.S. Thesis, Program in Media Arts and Sciences, MIT Media Lab.
33. Want R., Hopper A., Falcao V., and Gibbons J. (1992) The Active Badge Location System. *ACM Transactions on Information Systems*, vol. 10, pp. 99-102.