

# ReachMedia: On-the-move interaction with everyday objects

Assaf Feldman, Emmanuel Munguia Tapia, Sajid Sadi, Pattie Maes, Chris Schmandt  
*Ambient Intelligence Group, MIT Media Laboratory*  
20 Ames St., Cambridge, MA 02139  
{assaf, emunguia, sajid, pattie, geek}@media.mit.edu

## Abstract

*Mobile and wearable interfaces try to integrate digital information into our everyday experiences but usually require more attention than is appropriate and often fail to do so in a natural and socially acceptable way.*

*In this paper we present “ReachMedia,” a system for seamlessly providing just-in-time information for everyday objects, built around a wireless wristband with an RFID reader to detect objects that the user is interacting with. It enables hands- and eyes- free interaction with relevant information using a unique combination of audio output and gestural input, allowing socially acceptable, on-the-move interaction.*

*This paper will focus on the general merits of the design, and the novel gesture recognition technique used, showing 95% classification rate using the real-time online classification system.*

## 1. Introduction

With the growth of the Internet and the proliferation of digital information and services, many physical objects such as movies, books and other products now have extensive online information and powerful services associated with them, such as reviews, recommendations, and more. However, our interactions with this information almost always happen in front of a desktop computer using a mouse and a keyboard. We currently have no viable interfaces that allow us to combine the affordances and richness of our interaction with everyday objects, with the power and richness of the digital world.

Mobile devices such as smart phones have enough processing power and bandwidth to enable us to access these powerful services. However while this information may be accessible, it is not truly usable, mainly because the interaction modalities available in mobile phones are fairly limited. Current interfaces are

based on the desktop WIMP (Windows, Icons, Menus and Pointing device) paradigm, designed for interfaces utilizing the user’s full attention and hand-eye coordination, and not for interaction on the move. By “on the move interaction,” we mean not just interaction that is hands and eyes free, but also requiring minimal mental effort. Interaction on the move by definition happens while the user is engaged in some primary task other than the interaction itself, such as walking or flipping through a book. Therefore such systems should be unobtrusive and usable everywhere, not just in the home or within some special infrastructure.



**Figure 1: ReachMedia**

Commercial interfaces that support hands-free and eyes-free operation are using speech recognition currently, which still suffers from poor performance in noisy conditions. But as noted by Brewster [4], even if the performance of speech recognition technology were to improve, speech is not an ideal interface, because it is not socially acceptable. Not only do people feel embarrassed to use speech interfaces in public situations, speech is also intrusive to nearby individuals. This fact was indicated in a recent survey into the issue [22] and also by the growing phenomenon of cell-phone usage being banned in public locations such as buses and restaurants.

Wearable interfaces are also many times quite obtrusive. Jun Rekimoto [17] defined two socially

related requirements for designing wearable devices that are subtle and usable under various social contexts: first, they should support “hands-free” interaction in the sense that the device should not require the user to hold it in their hand, and even if the device requires the users to use their hands it should allow “quick changes between normal and operational mode”. Second, they should be socially acceptable. Rekimoto remarks that wearable interfaces to date do not usually address these design requirements. They are often rather obtrusive as is the case for the twiddler and glove-type input devices, and many times the fine-grained manual dexterity that these devices require demands that the user stand still to use the input device.

In our work, we are interested in interaction with physical objects such as books. As such, the interaction should not only be “hands-free” but also “eyes-free”. Our goal for this project is to design a system that gives users seamless access to the information and services related to physical objects in a way that compliments the natural affordances of our everyday interaction with these objects. Recent work [4] has shown the potential of gestural input and audio output for on the move interaction. In order to identify the subject of the interaction, we chose to use Radio Frequency Identifiers (RFIDs). RFID tags are passive chips with a unique identifier that are powered by inductive power transfer and capable of storing receiving and transmitting data. They are relatively cheap, unobtrusive, and do not require line of sight for reading.

This paper presents a novel system built around a wireless wristband that includes an RFID reader, 3-axis accelerometers and RF communication facilities. We view the wristband as a part of an on-body network that includes a cell phone and a wireless earpiece. The wristband allows the user an implicit, touch-based interaction with services related to objects by using RFID. Once an object is detected in the user’s hand, the user can interact with information about this object using slight wrist gestures, i.e. continuous movements of the hand in space. Thus, allowing an unobtrusive and socially acceptable interaction with everyday objects, yet a rich and engaging one.

## 2. Interaction design

One of the first design decisions we made was about the form factor of the device. We wanted the device to be simple and elegant, an accessory that we could imagine people wearing in public in everyday situations. Although a glove is a very effective form factor for both RFID [15] and gesture recognition [3],

we believe it to be too restrictive and inappropriate for everyday usage.

Advances in technology have significantly reduced the size and power consumption of all the components mentioned above, making it possible to consider small and elegant wearable devices. The wristband approach seemed more realistic especially because wrist-worn accessories, such as wrist-wallets and wristwatches are already a part of our common apparel. Given the current trend in miniaturization of electronic components, it is also not unreasonable to assume that the hardware would easily fit inside one of the above accessories within a few years.

In the next sections we will now detail our design consideration for the two major aspects of the interaction, namely the user interaction with the objects themselves and the user interaction with the information and services.

### 2.1 Touch based interaction with objects

In designing the interaction with the system, our primary challenge was to ensure that the affordances of the natural interaction with objects would be minimally affected as a result of the addition of the digital dimension. We therefore chose to base the interaction on the action of touching or holding an object, which also allowed us to make the detection process implicit, thus removing the intrusive process of explicitly having to ‘scan’ objects for interaction.

In choosing RFID as the tagging technology, we saw two main advantages for on-the-move interaction. First, unlike barcodes or infrared beacons, which require line of sight and explicit scanning, RFID tags are read using radio waves, thus allowing implicit event triggering. Users can concentrate on their primary task while the RFID reader detects the tag on the object they are holding or touching.

Second, the cost of the infrastructure required to support an RFID based system is low. RFID tags are battery-less and cheap. It is also expected that RFID tags will in the near future be used widely in retail, particularly for mid-range products such as books, because of their advantages over bar codes. Therefore in many cases the basic infrastructure incurs no extra cost.

Unfortunately, RFID technology also has its limitations. Primary among them is the reading range. We discuss the engineering aspects of this problem at length in the hardware section, but the interaction design aspects also merit discussion. Because the system is designed so that the detection of objects is implicit, the range of the reader is of great importance. If the range is too large, many events will be triggered,

and many of them will be “false positives.” For us, the form factor of the device has necessitated investigation into improving range, because if it is too short, the user will have to essentially scan the objects explicitly, and the detection process will require more attention than we consider desirable.

However, one of the primary design goals of this project is to provide an unobtrusive experience. Therefore, we chose to err on the side of caution in choosing a range which would limit the total number of “false positive” interruptions. Thus the interaction is “semi-implicit” meaning that we assume some type of conventions in the positioning of tags. On books for example tags might always be placed on the bottom of the spine and users will have to be aware of that if wishing to use the system. Such a convention will still allow the user to hold the book in a natural way and the reading will still be implicit but the user might need to be intentional especially if holding a very large book for example.

## 2.2 Gestural interaction with information

In choosing the mode whereby the user interacts with the system, a number of technical and human factors came into play. The final choice of input modality was motivated by social acceptability, intuitiveness, and separability. Although speech is a totally hands-free input modality i.e. doesn't require any usage of hands, as mentioned previously, we considered speech not to be ideal for input for our scenario since it has been found both by observation and research to be socially unacceptable in many public situations [22].

We chose to use gestures for this project because they can be made quite minimal and thus socially acceptable, while retaining ease of use. In choosing gestures for the system, we focused on gestures with as little as possible human-to-human discursive meaning. Head nods, for example, when performed in a public space might be misinterpreted by nearby individual. As such, we chose to focus on using minor wrist motions, which are outside the normal focus of attention. In considering sensors for this project, we found EMG, although extremely subtle, is still not robust enough and also somewhat invasive. We therefore chose to use accelerometers as input sensors for the wristband. We found rotations of the wrist to be minimal in terms of the amount of movement in space, yet highly recognizable by inertial sensors.



**Figure 3: The three gestures**

We chose a minimal set of commands, which we regarded as sufficient for supporting navigation of simple menus. The commands we chose to represent with gestures are: “Next,” “Previous” and “Select.” The gestural metaphors we used for these commands are slight flicks of the wrist. As an example, inward rotation of the wrist is “Next”, outward rotation of the wrist is “Previous” and a downward flick is “Select”. “Select” on a node causes the user to move down in the hierarchy. Moving up in the hierarchy is achieved by using the “previous” command when positioned on the first element in the hierarchy. Using “next” on the last element in the hierarchy causes the menu to wrap-around and move back to the first element in the hierarchy.

Our voice menu is a simplified version of the common Interactive Voice Response (IVR) type of menus, which uses speech labels and cues to guide the user through the hierarchy. Unlike IVR the system does not read the list of options when entering a new level in the hierarchy. Instead, each node just states the number of option it includes and the user can iteratively move inside the hierarchy using the “next” and “previous” commands. The leaves of the hierarchy are the results of the various services available for the object the user is interacting with. The context manager provides text to speech output for all nodes and leaves directly from the output of the services. Following the result of Pirhonen [16] we added non-voiced audio feedback to indicate the type of gesture detected as well as when an object is detected. Non-voice cues are also produced when the first or last items in the current hierarchy are reached.

An interaction with a book, for example, will take the following form: the user picks up the book, a slight notification sound indicates to her that the system has found services related to this book. The user flicks her wrist downwards, the short non-voice “select” cue is

played and immediately after a voice prompt says “ratings”, she flicks her wrist outward and hears the “next” command signal and the voice prompt says “reviews”; she signals “select” and the voice prompt says “New York Times”, she signals “select” again and listens to the review while flipping through the book’s pages.

### 3. Hardware

The hardware for this project is built on the MITes wireless sensor [13]. These modules use the nRF24E1 transceiver operating in the 2.4GHz Industrial, Scientific, and Medical usage (ISM) band, and include a microcontroller running at 16MHz. The MITes also include an external program memory, 3-axis ADXL202/210  $\pm 2g$  accelerometers and ancillary electronics for RF communication. The accelerometers data is sampled at a rate of 100Hz. According to the creators of the module, MITes are the smallest and lightest wireless 3-axis accelerometer sensors available to the research community [13].

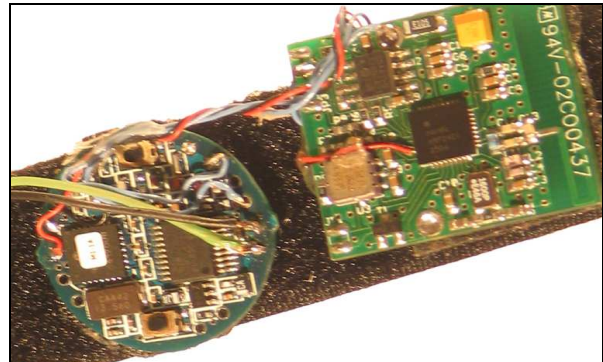
The RFID reader used is a SkyeTek M1 Mini, which has a diameter  $< 25\text{mm}$  and thickness  $< 2\text{mm}$ , which is optimal for the project requirements. The reader includes an integral 3V regulator, which powers the MITes, and is in turn powered by a 4.7V prismatic lithium polymer rechargeable battery with off-board charger for long-term use. The MITes microcontroller controls the reader via a serial connection using the SkyeTek m1-mini ASCII based protocol.

#### 3.1 RFID Technology

While RFID tags provide an excellent source of context information, they also pose problems. The tags used in this project are ISO15693 compliant 13.56MHz chip tags using inductive power transfer. Their availability and price are optimal for ubiquitous applications. However, they experience problems when bent (ie, when applied to the spine of a thin book). Previous industrial and academic similar systems, like Wan’s [20] ‘*Magic Medicine Cabinet*’ or smart bookshelves, use RFID readers that are embedded into static objects. This makes it possible to use fixed reader antennas, which usually does not exhibit these pathologies due to power availability and the more extensive drive electronics.

Unfortunately, the 2.3cm diameter of the integral antenna, combined with low drive current for the outgoing carrier causes the effective range of the standalone reader to as small as 4 to 5cm. To increase the range, the wristband form factor was used to

increase the antenna diameter to the diameter of the wrist. However, the reader by itself failed to produce sufficient output to drive this antenna. To that end, an RF-mode amplifier was designed to increase the output of the reader. This has doubled the range and to approximately 10 cm of range. While still a 3cm shorter than our goal of 13cm, it is sufficient to allow natural interaction with objects such as book. We are currently investigating methods of separating the input and output waveforms and bi-directional amplification methods, in order to increase the range further.



**Figure 5: ReachBand including the MITes board (right) and the Skyetek m1-mini RFID reader (left)**

### 4. Context Resolution Service

Mobile context-aware devices pose an interesting set of software design choices. Due to the limitations of the mobile interaction platform, it is usually necessary to handle the management and processing of complex or resource-intensive context information off-board, and the need for “context services” for such applications has been long recognized [1,6]. While there has been some dissent from those interested in automation, in general, an off-board processing model provides many benefits in terms of mobile node and communication protocol complexity, and is the model used in the Context Manager, the object resolution and service management framework designed by our research group and used in the ReachMedia project.

The Context Manager has a threefold task of converting RFID identifiers into object metadata and context, accessing services with the available context and session information, and transferring the context information back to the device for display. This is done with “connectors” that translate the incoming RFID sensor data into internal session data, and then converting the output of the context-relevant services for the output modality. By separating out object resolution, it is possible to minimize wristband data

bandwidth, while reaching a compromise that allows many services to use the data [5]. This has the additional effect of reducing the demands on the RFID tags, which can now contain minimal metadata. On the output side, the output connector capitalizes local processing to provide optimally translated output, thus reducing overhead and requirements for the target mobile device.

## 5. Gesture recognition

There are two approaches to the dynamic gesture recognition problem [12]: Discrete Gesture Recognition (DSR) and Continuous Gesture Recognition (CGR). DSR uses an explicit command from the user to indicate the start and stop of a gesture via a non-gestural modality, e.g., with a button. With CGR, on the other hand, the recognition is carried out online from a continuous flow of gestural input data.

Naturally, implementing a DSR system requires the user to hold a secondary device, which conflicts with our hands-free design choice. The system therefore was designed as a CGR, which posed two major challenges: first, detecting an activity in the signal flow and second, correctly identifying noise or “unknown” gestures. The latter is especially problematic in a system such as ours because the “unknown class” includes “everything,” while at the same time the system requires a low false-positive rate.

Much of the research involving gesture recognition systems [23,12], uses Hidden Markov Models (HMMs) for classification due to the signal’s temporal character. The input to the model is usually the raw signal vector, or a scalar resulting from vector quantization (over the axis vector) [12]. However, the problem we wish to solve is subtly different from this general case. First, the gestures we are proposing are very slight and short, and second because we are using CGR, we need to filter out noise with a very low false-positive rate.

After examining a variety of approaches, we chose an algorithm that uses a custom set of time domain features and a discretized Naïve Bayes classifier that has shown equivalent results to the HMM approach (see preliminary test section). The advantages of the naïve Bayes classifier are computational efficiency, simple implementation and reliability.

### 5.1 Activity Detection

Our system targets mobile devices by design, and thus efficiency is a necessity. Therefore, instead of constantly classifying every part of the signal, the

system uses a preprocessing stage to detect activity in the signal and classifies only the parts where activity is present.

We detect activity by using a simple axis-by-axis variance window with preset thresholds using the methodology introduced by Benbasat [2], whose work also focused on mobility. He has shown that the per-axis variance window method is both very effective and is easily optimized for real-time classification. Benbasat’s work also indicated optimal start and stop thresholds values as a function of the window size. We have chosen to use a 16-sample window with a start threshold of 100 and a stop threshold of 50. When the variance of one or more axis passes the “start” variance threshold, we begin recording the gesture. We stop recording once the magnitudes of all of the axes are below the “stop” variance threshold. Naturally, this method required the user to have a short pause before and after the gesture. At the current sampling rate, this pause is <100ms per gesture, and generally imperceptible in normal usage.

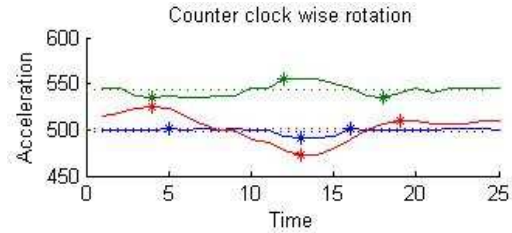


Figure 6: Acceleration signals samples of the ‘next’ gesture with the inflections as \*

### 5.2 Feature extraction

For each gesture with  $n$  samples, where each sample is a vector of length 3, and each element in it represents an axis, we extract the following 25 features:

**Length (1)** – the number of raw signal vector samples included in the gesture,  $n$  in this case.

**Power (2-4)** – the energy in each axis is calculated according to Parseval’s theorem as follows:

$$P(\bar{a}) = \frac{1}{n} \sum_{i=0}^{n-1} a_i^2$$

**Cross Correlation (5-7)** – The pairwise similarity of the signal on two different axes, as measured using the correlation coefficients that are calculated for a given gesture as:

$$r(A, B) = \frac{n \sum a_i b_i - \sum a_i \sum b_i}{\sqrt{n \sum a_i^2 - (\sum a_i)^2} \sqrt{n \sum b_i^2 - (\sum b_i)^2}}$$

We calculate  $r$  for all the pairwise combinations of axes, in our case:  $xy, xz, yz$ .

**Inflections (8-25)** – The rest of the features are a result of a signal-processing algorithm designed to find the 3 most significant inflections in the signal. The choice of 3 points arises from the observation that our set of gestures has three main velocity change points: start, turn back and stop.

Finding the significant inflections requires us to have a reference by which to measure the magnitude of the signal relative to its starting point i.e. a zero point. The algorithm utilizes some special properties of the signal to do so. Because the gestures start and end at approximately the same point in space and with zero velocity, the static acceleration due to gravitational effects on each axis at the beginning and end of the gesture should be at least approximately equal. Furthermore, the mean acceleration throughout the gesture should also approximately be equal to the static acceleration value. Therefore, we can use the mean as a reference point for calculating a “relative magnitude” of a sample, which is the difference between the sample value and the gesture mean. The magnitude of the sample is then the absolute value of the relative magnitude. We then choose the three maximal peaks from both sides of the mean (figure 4).

### 5.3 Preliminary tests

We compared the performance of the described time-domain feature based approach with the raw data HMM approach. We performed these tests offline on data collected from one user with 197 gestures, with approximately 60 from each type and 96 samples of noise data, which was collected by recording few typical user’s daily activities and interactions.

We used a left-to-right HMM model and used 10 fold cross validation to find the optimal number of states for the gestures model, which we found to have 3 states for each gesture and 5 states for the noise class. The results showed that the discretized Naïve Bayes performed as well as the HMM, with 98% accuracy on the data with 20% of the samples used for testing.

Although these preliminary results used only one subject, we found them sufficient evidence to indicate that the Naïve Bayes was a comparable algorithm and was a better choice for the system given its efficiency, especially due to our assumption of a relative low variance in gestures across users. Also, this assumption was proven to be correct in the user-independent versus user-dependent tests, which we discuss in the next paragraph.

### 5.4 User study

In order to evaluate the accuracy of the naïve Bayes online gesture recognition system, we trained and tested the system with 10 subjects: 8 males and 2 females, all colleagues. Each subject trained the system for 2 minutes using 60 gestures (approximately 20 samples of each gesture, selected at random by the system). The subjects then tested the resulting model with another set of 30 gestures. The training and testing were done by an automatic system that randomly prompted the users to perform a gesture every 2 seconds.

<b>Next</b>	0.97	0.01	0.02
<b>Previous</b>	0.04	0.96	0
<b>Select</b>	0.02	0.02	0.96

**Figure 7: Between-gesture confusion matrix for 779 gestures samples**

The results were very encouraging with 4 users having 100% accuracy; and an average accuracy of 94.8% with a variance of 30%.

Since the results above are achieved by using user-dependent models we wished to evaluate what the accuracy cost would be when using a user-independent model. From a user experience point of view, the system is relying on a personal device. Hence, a first time, short training session that takes less than 2 minutes, does not seem unreasonable. Nonetheless it is important to get a measurement of how much more accurate a user dependent model is in order to assess its necessity. In order to get this measurement we used a 5 fold cross validation analysis using all our data. For each fold we trained a model using eight subjects and tested it on the remaining two. The resulting average accuracy was 91.2% with a variance of 10.1%. We consider these results to be high enough to justify further evaluation of the usability of the user-independent gesture recognition system.

### 6. Related work

There exists a rich body of work, which uses sound, head movement, body gesture, and augmented/tagged objects to trigger and manipulate interaction. Early work is the field used cameras, color tags and visual displays to create an augmented interactive reality. Such are Nagao’s [14] “Ubiquitous Talker” that used speech and Starner’s [19] “Physically based hypertext” where stepping toward an object indicates hyperlink clicking. Very recent work- called ShopAssist [22]

augmented the physical activity of shopping in an electronic shop and focuses on examining few input modalities. Although ShopAssist does not use mobile RFID input or gestures the way we defined them, a very applicable result of that project to our system is their finding that while speech was a preferred modality in laboratory tests it was found, in their test field, unsuitable for public scenarios.

We will review other related work separately while focusing on the individual technologies that together make up our system.

## 6.1 Multimodal interfaces for on-the-move interaction

The widespread usage of mobile phones and the realization of the limits of their input and output modes, i.e. their small screens and keyboards, highlights the need for new interaction methods for mobile devices. This realization has given rise to extensive research in the field of multimodal interfaces to enhance the usability of mobile devices and to enable a new level of services and information for users that are on the move.

Some of the key research that our work builds on is Brewster's investigation of the combination of gesture input and audio output for interactions on the move [4]. In this work two systems were presented and evaluated; both used speech and non-speech audio output in combination with gesture input. The first system was a 3D audio radial pie menu that used head gestures for selecting items. The second system was a sonically enhanced 2D gesture recognition system for use with a belt-mounted PDA. Their results show that sound and gesture can significantly improve the usability of a wearable device in particular under 'eyes-free', mobile conditions.

The use of voice and audio interface for mobile systems is not new. Sawhney and Schmandt [18] previously presented a 3D spatialized audio system, "Nomadic Radio," which used voice and non-voice input and output methods. The usage of gesture in a wearable context was also explored and discussed in Rekimoto's GestureWrist [17]. In this work, a wristband with capacitive sensors was used to recognize hand gestures and forearm movements, focusing in particular on minimally obtrusive interaction. Preliminary work investigating subtle and intimate muscle movements and EMG has also shown great potential for on-the-move interaction [7]. However, most of the work in gesture recognition relies on interfaces that we consider unsuitable for the types of applications we are interested in. Many systems use a glove interface, such as in [3] which in

our mind is too constrictive an interface for usage in everyday life.

In recent years accelerometers are gaining popularity as mobile sensors due to their low price, low power consumption and robustness. However, comparatively little work has been done applying accelerometers to the gesture recognition problem. In particular, Benbasat [2] has created a framework for rapid development of discrete gesture recognition applications. Also, Mäntyjärvi [12] has used mobile phones, enhanced with accelerometers, to allow the use of gesture control, providing some highly relevant insights into this topic. Hinckley [9] has used static 2D acceleration signals (tilt sensors) to allow background sensing of user actions and in turn to allow a device to take proactive action. Lastly, body worn accelerometers were also used in research on activity detection and recognition [11].

## 6.2 RFID augmented objects

Using RFID for mobile interaction was suggested by Hull [10] for tagging places and people, Want [21] further explored the technology and suggested using it to enable interaction with physical objects, which has since become a popular line of research [20, 8] in the pervasive computing community.

Typically, mobile pervasive systems that use RFID for augmenting everyday objects have relied on explicit (i.e. intentional) actions of the user. Often this design decision has been forced by the fact that the input and output devices used in these projects were PDAs that were too big to support implicit input. However, recent work at Intel Research Labs has looked into activity detection for health care application [15] using an RFID reader integrated into a glove.

## 7. Future work and discussion

In this paper, we present a holistic approach to the problem of on the move interaction with augmented objects, integrating solutions from user input, to processing of input, to user output. On the move interactions present a completely new class of interaction issues, and thus require a comprehensive approach. Our work thus far has been focused on the technical implementation and development of the hardware, gesture recognition software, RFID antenna and context manager. We now have a working system with reasonable RFID range and very reliable gesture recognition. As a short-term goal, we intend to do more extensive user studies to evaluate a range of

issues. In particular, we are interested in whether the interface truly satisfies the definition of “on-the-move interaction” as discussed in the introduction.

Also, from personal experience with the system we know that the usage of slight gesture raises issues of usability, namely the ability to remember the exact way gestures were performed during training. We believe there is much more work to be done in order to get a better understanding of the usability issues and consideration in designing a system using such gestures.

## 8. References

- [1] G. Abowd, C. Atkeson, J. Hong, S. Long, R. Kooper, M. Pinkerton, “Cyberguide: A mobile context-aware tour guide”. *ACM Wireless Networks*, 3:421--433, 1997.
- [2] A.Y. Benbasat and J.A. Paradiso “Inertial Measurement Framework.” *International Gesture Workshop, GW 2001*, London, UK
- [3] T. Baudel and M. Beaudouin-Lafon. “Charade: remote control of objects using free-hand gestures”. *Communication of the ACM*, 36(7): 28–35, July 1993.
- [4] S.A Brewster, J. Lumsden, M. Bell, M. Hall and S. Tasker “Multimodal 'Eyes-Free' Interaction Techniques for Wearable Devices”. *Proc. CHI 2003*.
- [5] G. Chen and D. Kotz. “Context aggregation and dissemination in ubiquitous computing systems”. *Workshop on Mobile Computing Systems and Applications*. IEEE Computer Society Press, June 2002.
- [6] N. Cohen, A. Purakayastha, J. Turek,, L. Wong, and D. Yeh “Challenges in flexible aggregation of pervasive data”. IBM Research Division, Thomas J. Watson Research Center. Technical Report RC21942.
- [7] E. Costanza, S. A.Inverso, R. Allen. “Toward Subtle Intimate Interfaces for Mobile Devices Using an EMG Controller”. *Proc. CHI 2005*, April 2005.
- [8] C. Florkemeier, M. Lampe “Issues with RFID usage in ubiquitous computing applications”. *Proc. Pervasive 2004*.
- [9] K. Hinckley, J. Pierce, E. Horvitz, and M. Sinclair. “Foreground and Background Interaction with Sensor-Enhanced Mobile Devices”. *ACM Transactions on Computer-Human Interaction (TOCHI) 2005*, 12(1), 31-52.
- [10] R. Hull, P. Neaves, and J. Bedford-Roberts “Towards Situated Computing” *Proc. ISWC 1997*
- [11] P. Lukowicz, J.A. Ward, H. Junker, M. Stäger, G. Tröster, A. Atrash, T. Starner: “Recognizing Workshop Activity Using Body Worn Microphones and Accelerometers”. *Pervasive 2004*.
- [12] J. Mäntyjärvi , J. Kela, P. Korpipää, S. Kallio, “Enabling Fast and Effortless Customisation in Accelerometer Based Gesture Interaction ” *Mobile and Ubiquitous Multimedia*, Maryland, USA, Oct. 27, 2004.
- [13] E. Munguia Tapia, N. Marmasse, S. S. Intille, and K. Larson, "MITes: Wireless portable sensors for studying behavior," *UbiComp 2004*.
- [14] K. Nagao and J. Rekimoto, "Ubiquitous Talker: Spoken Language Interaction with Real World Objects" *Proc. of Inter. Joint Conf. on Artificial Intelligence (IJCAI) 1995*
- [15] M. Philipose, K. Fishkin, D. Fox, H. Kautz, D.J. Patterson, M. Perkowitz. “Guide: Towards Understanding Daily Life via Auto-Identification and Statistical Analysis”. *UbiHealth 2003*.
- [16] A. Pirhonen, S. Brewster, and C. Holguin. “Gestural and audio metaphors as a means of control for mobile devices”. *CHI, 2002*. Minneapolis, Minnesota, USA
- [17] J. Rekimoto “GestureWrist and GesturePad: Unobtrusive wearable interaction devices”, *Proc. ISWC 2001*, 21—27
- [18] N. Sawhney and C. Schmandt, "Nomadic Radio: Speech & Audio Interaction for Contextual Messaging in Nomadic Environments," *ACM Transactions on CHI*, vol. 7, pp. 353-383, 2000
- [19] T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healey, D. Kirsch, R. Picard and A. Pentland, "Augmented Reality Through Wearable Computing", *Presence* 386-398 Winter 1997
- [20] D. Wan "Magic Medicine Cabinet: A Situated Portal for Consumer Healthcare" in *Proceedings of First International Symposium on Handheld and Ubiquitous Computing (HUC '99)*, September 1999.
- [21] R. Want, K.P. Fishkin, A. Gujar, and B. L. Harrison “Bridging Physical and Virtual Worlds with Electronic Tags”. In *Proc. of CHI 1999*
- [22] R. Wasinger, A. Krüger. “Integrating Intra and Extra Gestures into a Mobile and Multimodal Shopping Assistant”. *Pervasive, 2005*.
- [23] T. Westeyn, H. Brashear, A. Atrash, and T. Starner. "Georgia tech gesture toolkit: supporting experiments in gesture recognition." *International Conference on Multimodal Interfaces (ICMI)*, 2003.