HAUPTSEMINAR MEDIENTECHNIK

„SENSOR-BASED USER INTERFACES – SCIENCE OR SCIENCE FICTION?“

WINTERSEMESTER 2009/2010

Technische Universität München

Lehrstuhl für Medientechnik

Prof. Dr.-Ing. Matthias Kranz

SENSOR-BASED USER INTERFACES

Advances in computing technology enable novel and innovative user interfaces, formerly only known in science fiction movies. Research on sensors and sensing systems, miniaturization and increase in processing power and novel approaches to human-computer interaction led to the development of unconventional sensor-based user interfaces.

In the seminar 'Sensor-Based User Interfaces', in the context of the Hauptseminar Media Technology, a selection of enabling sensing technologies, algorithms and resulting user interfaces, from research and science fiction, was investigated. Several recent science fiction movies and their user interfaces were discussed regarding their potential and practical realizations. This includes multitouch, interaction across multiple surfaces, gesture-based interaction, tangible user interfaces, augmented reality, mixed physical and digital interaction devices, and many more. Application areas include data management, gaming and interaction with and in intelligent environments.

The thesis collected in this technical report give insights into 6 selected topics:

- Acceleration Sensor-based User Interfaces
- Tangible User Interfaces
- Smart Surfaces
- Capacitive Sensing based User Interfaces
- Game User Interfaces
- User Interfaces for Media and Creative User

We thank all students for the great work and their participation!

The LMT Team:.

Fernanda Brandi
Michael Eichhorn
Julius Kammerl
Werner Maier
Yang Peng
Martin Pfannenstein
Fan Zhang

Prof. Dr.-Ing. Matthias Kranz

TABLE OF CONTENTS

Acceleration Sensor-Based User Interfaces

Peter Christian Abeling
peter.abeling@mytum.de

January 7, 2010

Abstract

Today acceleration data is used in many different types of user interfaces. This paper gives an overview of different applications and data processing methods which are commonly suggested in current research and enrich ubiquitous computing with additional possibilities of interaction. Several data classification methods are also reviewed, as they are an important part in sensor-based user interfaces and open a wide range of usage. The focus lies on tangible objects, gesture and human pose recognition as well as authentication methods. We will also give the reader an insight into sensor hardware and communications and therefore get a full understanding of the acceleration sensor's potential.

Many of the introduced problems in this paper are common for pattern recognition tasks. However, here they are explained from the viewpoint of acceleration sensor-based user interfaces. This allows a detailed look at modern techniques for processing acceleration data.

1 Introduction

Ubiquitous computing is and becomes part of our everyday life in different environments and objects[1]. As everyone should have the chance to use and control complex technical systems for different purposes like learning, working or entertainment, the user interface holds a central position. It should be as easy and intuitive as possible to include a wide range of possible users. Acceleration sensors are a well established possibility to develop user interfaces for motion control and recognition. Every motion causes acceleration, no matter if an object is moved or the human body moves itself.

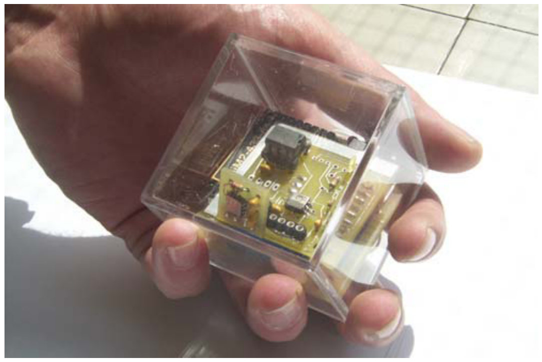
The success of using acceleration data for recognising gestures can be seen in Nintendo's *wii*, which is a popular product on the entertainment market. But acceleration can not only be used for motion game control. Other applications are also researched like intuitive multimedia control, activity recognition and authentication. These applications are reviewed here at first in order to give the reader a short overview of the status quo. This is followed by some aspects of sensor hardware and acceleration data, as well as explanations about the functionality of two different sensor types.

The last section concentrates on different classification methods for acceleration data. For giving the reader a reference, different examples are taken to explain the choice of a classifier or feature extractor. Nevertheless, this paper does not cover all possible classifiers, as there are a big number of methods and many modifications. The focus rather lies on methods which are researched in representative papers in order to get the reader an insight into acceleration sensor-based user interfaces.

2 Overview of Applications

There is a wide range of applications using acceleration sensor-based user interfaces. All different applications are based on gesture recognition methods, therefore the main aspects of modern gesture recognition are explained in this section. Afterwards different applications are distinguished, which all focus on different possible fields of usage. This contains intuitive multimedia control, activity recognition and authentication applications.

[1] This fact is already indicated by the term "ubiquitous".



(a) A cube as a tangible user interface with acceleration sensors (from [4]). Due to their geometry, cubes have proven to be intuitive objects.

(b) A test user wears acceleration sensors for recognizing basic activities like climbing stairs. In this example many sensors are used as a *Spine* (from [5]).

Figure 1: Examples for acceleration sensors in different applications. The cube as a tangible user interface (*a*) and wearable acceleration sensors for gesture recognition (*b*).

2.1 Gesture Recognition

Gesture recognition, as a part of human-machine interaction, tries to interpret human gestures with mathematical models and methods. In the presented applications the sensor data is provided by acceleration sensors which are positioned at either the human body itself (including its *wearables*) or tangible objects which are physically moved by a human (see section 3). In general, predefined gestures are trained by the user or learned autonomously in order to recognize them, while using a device.

An example for classic gesture recognition is the experiment which uses a wearable acceleration sensor to recognize Thai Chi movements presented by Kunze et al. [1]. Athletic sports movements can be complex and difficult to recognize, but the data could be used to improve training methods or fine-tune the quality of movements. The authors showed that it is possible to even distinguish between different levels of expertise with only using eight sensors spread over the body.

Gesture recognition is also used in intuitive game control. A famous example is Nintendo's *wiimote*. In [2] Heinz et al. research wearable sensors in their potential to recognize martial arts movements in real-time in addition to other sensors[2]. These applications lead to intensive user interaction and a new state of gaming in general. A comparison of motion based gaming with acceleration data is given by Champy [3]. The author also considers the impact which the development has on game design in the future. It means, that motion control leads to different game genres in which physical elements and motion hold a central position. This has to be considered from the beginning of a game design workflow and even leads to completely new game ideas.

2.2 Intuitive Multimedia Control

The focus of other research lies on the intuitive aspect of user interfaces, especially in multimedia control applications. Many experiments use a cube as a tangible object, as it has many possible applications and its position can easily be tracked by the help of only a few acceleration sensors.

The cube can be used for general input and navigation exercises (Van Laerhoven et al. [4], Foxlin [6]). A specific example is a user interface for home entertainment systems prepared by Block et al. [7]. Moreover, the cube could be used for playful learning applications for children as shown by Terrenghi et al. [8] which fit to the intuitive character

[2] In this case gyroscopes are also used.

of the cube. It could also be used for showing digital contents like News, Webpages or Videos etc. in a modern way (Matsumoto et al. [9]). An example for such a cube is given in Figure 1(a).

Other devices which use acceleration data, like the *i-Throw* presented by Lee et al. [10], can be used to control ubiquitous environments. Again gestures are used to prepare user-friendly interaction with different other devices like displays, speakers or different tangible objects.

2.3 Activity Recognition and Detection

Recognizing the activity or pose of a human is a difficult task. Foxlin uses acceleration sensors as an addition to improve other techniques [6]. The acceleration data measured in the shoes contain valuable information about the orientation of the person. Another advantage is that acceleration sensors are small and cheap (see 3.1).

An example of an experiment setup for activity recognition can be seen in Figure 1(b).

2.4 Authentication

Studies about the reliability of acceleration sensor data for authentication appliances like from Okumura et al. [11] lead to another different field of possible implementations. It is shown that a human's individual shaking of a device, like a cell phone, can be used for verifying the correct owner, instead of using a password or other biometric data. Another field of usage is the device-to-device authentication researched by Mayrhofer et al. [12]. Especially wireless devices can be paired and shaken together in order to verify a secure connection. An example for such an application is a mobile phone and headset.

3 Acceleration sensors and data

In the previous section an overview of different applications is presented. The data which is used to recognize gestures is produced by acceleration sensors. This section concentrates about different types of sensors as well as on their functionality. Additionally, a short overview about communications is given, as data often has to be transferred from a mobile device to a terminal.

3.1 Aspects of Sensor-Hardware

Sensors used for user interfaces have to fulfil some expectances due to their size, price, and functionality. Modern sensors can be purchased for about 5-8 USD and can be produced very small (see [4],[5]) which allows them to be integrated in small objects like the presented cube (see section 2.2), mobile devices, game controllers, etc..

Sensors produce analogue data which has to be sampled in order to further data computation. For example, this can be done with a sampling rate from $100 - 600$ Hz for appropriate results [12]. There are sensors which can have a measure range up to $\pm 10g$ [11]. Many of the applications presented additionally use a microprocessor for preprocessing the data and an A/D-Converter. The data can be smoothed to avoid inaccuracies and therefore to get more reliable results. The caused delay can be accepted in many cases.

3.2 Functionality

Different sensors use the same idea of measuring acceleration and gravity. A sensor can be modelled as a tiny ball which is adjusted by two springs on both sides. The ball is limited in its movement by a small tube (see Figure 2 a)). The ball's position is measured to get output data capacitively. This enables the sensor to measure static acceleration (gravity) as well as dynamic acceleration in two dimensions (e.g. X-Y). The discrimination of static and dynamic data was researched by Veltink et al. [13]. Using high- and lowpass filter, rectification and a threshold, the signal can be recognized as recent movement, which in this case is equal to dynamic acceleration.

Purchasable sensors use two orthogonal positioned base-sensors to gain 3-dimensional data. The redundant data (e.g. the double covered X-axis) can be used to improve data quality. Figure 3 shows the physical arrangement and the abstract model inside a cube as an application example. The principle can be assigned to every other application with accelerometers which gain three-dimensional data. In this case, the acceleration data is normalized in every dimension in order to get values in the interval $[-1.0, 1.0]$.

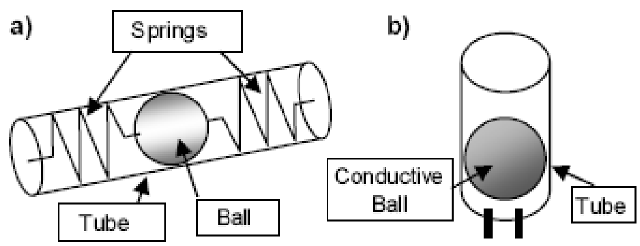


Figure 2: Abstract models for two different sensor types: a) shows an accelerometer with a ball adjusted by two springs inside a small cylinder. b) shows the ball switch which is a binary sensor. From [5]

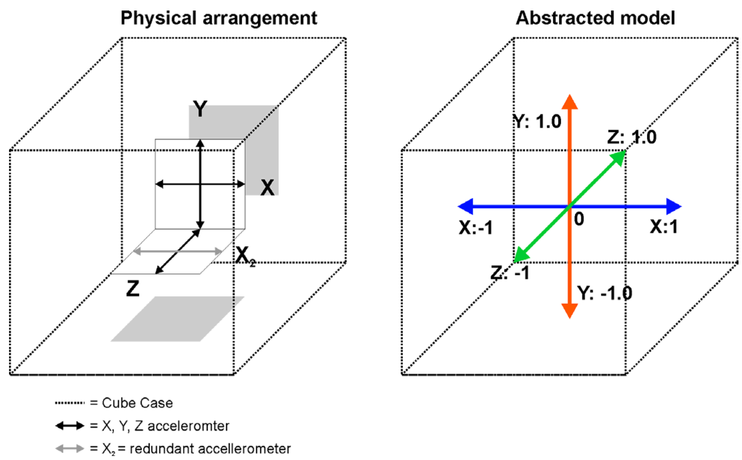


Figure 3: Physical arrangement and abstract model of two sensor devices inside a cube. The left image shows two orthogonal adjusted accelerometers. The right image shows the appropriate abstracted model. From [7]

Another method of measuring acceleration data is reviewed by Van Laerhoven et al. [5]. The authors use many ball switches to combine them to a *porcupine*. This low power method is investigated due to its potential to compete against a classic acceleration sensor. The ball switch contains a roller ball which is conductive and can close a switch inside a small tube (see Figure 2 b)). The binary sensors are placed in different directions to get three dimensional information.

3.3 Communications

As the gathered acceleration data often is gained in mobile devices, it has to be transferred to a terminal for reasons of further processing. In fact, communications electronics is the most power consuming part. As the data has to be transferred over short ranges, frequency modulation and a data rate of about 64kbps is used (see Van Laerhoven et al. [4]). Other possibilities are Bluetooth or ZigBee like used by Lee et al. [10], as they are optimized for short range communication.

4 Classification of Sensor data

Classification plays an important part in every acceleration sensor-based user interface. It can be seen as the interpretation of sensor data in order to initiate an event or to simply recognize a gesture. The main problem is the fact that users perform gestures differently in time and space. Even a repetition of the same gesture by the same person will be different, but has to be recognised as the same. Moreover, the accuracy should be high to not unsettle possible users.

Dependent on the application the user should have the possibility to train his or her own gestures with few repetitions. This again demands a well-fitting model and good convergence. Other difficulties lie in the complexity of the algorithms, because they often have to be calculated in real-time.

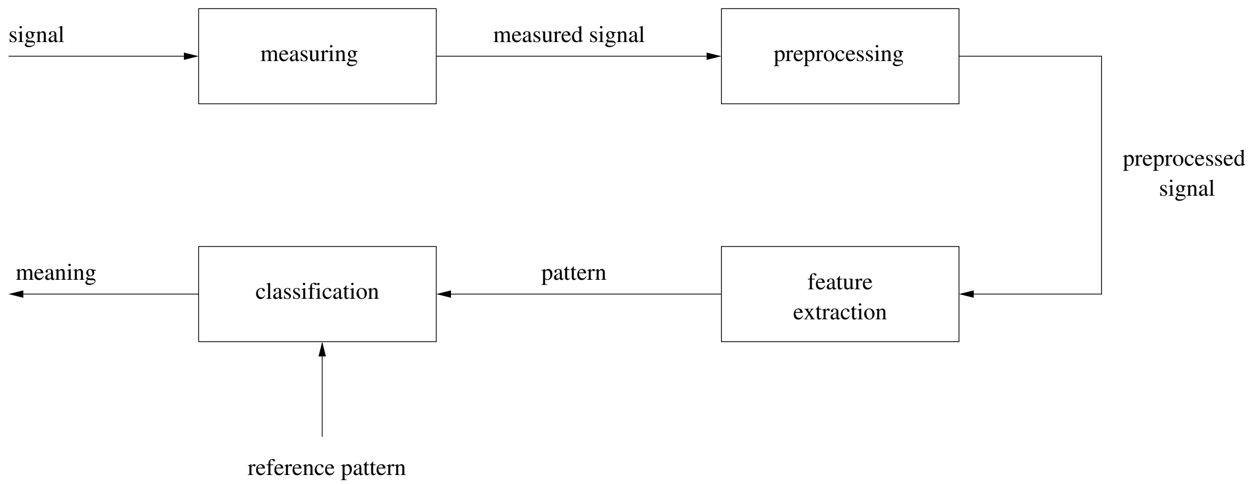


Figure 4: Pattern recognition system model. After having measured and preprocessed the signal, features are extracted in order to describe the signal with few parameters. These parameters are called patterns. In the last step the pattern is classified with the help of reference patterns in order to get the meaning of the signal. The reference patterns are gained by training data.

Basically, the classification of gestures is one step of a pattern recognition system. A classic system is shown in Figure 4. After having measured the input signal, it has to be preprocessed. The edited signal is afterwards used to extract features which should be unique, deterministic and decorrelated for the gestures they describe. Finally, the features are used to classify the pattern with the help of reference samples.

In this section some popular methods are reviewed. As certain approaches fit better than others in given contexts, the classification methods are connected to an example. The preprocessing and the extraction of features for a certain set of data are reviewed, as well as classification methods like Maximum likelihood estimation, Hidden Markov Models (HMM) and neural networks.

4.1 Preprocessing

Before classification is done, the sampled acceleration data has to be preprocessed. Assuming that the data is three-dimensional (due to three acceleration directions), it is normalized to equal length and amplitude (see Mäntyjärvi et al. [14]) in every time step. If equal vectors are needed, inter- and extrapolation can be used. Mäntyjärvi et al. use the data to train a HMM (see 4.5) which requires one-dimensional data in this case. Therefore, vector quantisation is used to reduce the dimension to one.

4.2 Feature extraction

In order to be able to compare different signals, it is a common approach to extract features which should describe the signal within a specific period of time explicitly. Common features are statistical values like variance or correlation in the time domain, or Fourier-components in the frequency domain. An example is given by Mayrhofer et al. [12]. For comparing two devices shaken together, the authors name two major problems. The first problem is the temporal asynchronism. In order to compare the corresponding values, triggering has to be used. The other more difficult problem is the spatial asynchronism. According to how the user holds both devices, the sensor's dimensions are not aligned. Therefore the authors decided to only take the magnitude of all three dimensions to get a one-dimensional direction-independent signal. For assuring robustness against small variations, quantised FFT features are used as features for every time interval.

Heinz et al. focus on the explosiveness of martial arts movements [2]. Features are calculated by using frequency range power coefficients (FRP). For every frequency band the power of the FFT-components is calculated which represents the explosiveness more precisely than FFT components. Another approach is the entropy of frequency in a windowed time interval. Visualising these entropies leads to clustered structures which represent different movements. Afterwards linear classification algorithms (see section 4.3) are applied to classify new data.

For identifying users carrying a portable device in [15] Mäntyjärvi et al. use other statistical features. The new data to be classified is cross-correlated with templates of training data. This can be done, because the shape of the acceleration signal is assumed to be unique for every person. The cross-correlation is therefore a measure for the similarity of the two signals.

4.3 Linear classifier

There are several linear and piecewise linear classifier which are used to solve linear classification problems, like the *Nearest-Mean-Classifier* or the *k-Nearest-Neighbour-Classifier*. The *Nearest-Mean-Classifier* uses the training data's mean values of the different classes, in order to calculate separation lines between them. The *k-Nearest-Neighbour-Classifier* at first calculates the Euclidean distance from the new data to all training data. Afterwards the classification is done by a majority decision of the k nearest training data values.

As classification of acceleration data is only in exceptions a linear problem, other, more complex classifiers will be introduced in the next sections.

4.4 Gaussian modelling and Maximum Likelihood estimation

As explained in 2.2, the cube is an intuitive user interface due to its geometry. In order to recognize which side is up, a four-dimensional Gaussian $G_i(\overline{x})$ is used to model the cube:

$$G_i(\overline{x}) = \frac{1}{(2\pi)^2\sqrt{|\Sigma_i|}} e^{-\frac{1}{2}(\overline{x}-\overline{\mu}_i)^T \Sigma_i^{-1}(\overline{x}-\overline{\mu}_i)}. \qquad (1)$$

Van Laerhoven et al. model every side with a Gaussian [4]. i ranges from 1 to 6 for every side of the cube, $\overline{x}$ represents the four-dimensional vector for the four sensor channels. μ is the mean average vector and Σ the covariance matrix.

Furthermore, the cube can be modelled as a finite state machine with a distinct set of states (for a cube there are 6 states). The side, which is on top, can now be calculated with a maximum likelihood estimation. The state the cube has, is the most likely state concerning the Gaussians. The Gaussian modelling belongs to the class of static-nonlinear classifiers.

4.5 Hidden-Markov-Models

Hidden-Markov-Models belong to the stochastic signal models. In [14] Mäntyjärvi et al. explain the functionality of HMMs concerning their usage for gesture training and recognition. The idea of modelling time-series with spatial and temporal variability with HMMs is successfully used in speech recognition and video-based gesture recognition. The aim is that the user can train his or her own free form gestures with a minimum of repetitions. The input signal can be a one-dimensional sequence of discrete symbols, an example are the feature vectors explained in section 4.2.

The Hidden-Markov-Model is based on two random processes. The first is modelled by a regular Markov chain described by states and transition probabilities. These states cannot be observed from outside, they are hidden. The second random process produces observable output symbols according to state-dependent probability distributions. The task is to conclude from the observed symbols to the hidden states in order to reconstruct the parameters of the HMM, which lead to the observations with the highest probability.

Taking the functionality of a HMM into account, the recognition of a gesture is done by finding the model which produces the maximum probability of the observed symbol sequence. Therefore for every gesture an own HMM is calculated and its parameters are learned from the training data.

4.6 Neural networks and Self-Organizing Maps

The last approach which is presented in this paper to classify the acceleration data is used in several experiments like by Randell et al. [16], Baier et al. [17] and Van Laerhoven et al. [5]. Neural networks are a clustering algorithm to classify multidimensional input data in strong correlation to the functionality of the human brain.

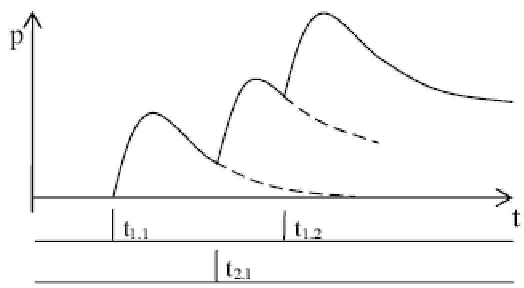


Figure 5: Example for spiking neural networks. Each time a spike arrives, e.g. at $t_{1.1}$, the action potential p increases. If a certain threshold is met, it "fires". As seen in this figure, the potential grows for every of the two input lines. From [5]

The method which is appropriate for gesture recognition from acceleration data is called SOM (*Self-Organizing-Map*) and is amongst others presented by Baier et al.. SOMs consist of two different layers, the input layer and the map layer. The map layer is two-dimensional (in this application) and can be seen as a mesh grid of neurons. Every neuron of the map layer is connected to every input neuron, as well as to every other neuron in the map layer. Furthermore, every connection from the input layer to the map layer has a weight ω_i. During the training phase all euclidean distances between the input value and the weights ω_i are calculated in order to assign the best matching neuron which has the smallest distance. Afterwards the impact on the best matching neuron and neighbourly neurons of the map layer is calculated (e.g. with a Gaussian) and their weights are updated. The whole training is repeated with different training samples in order to achieve a good organisation of the map. New input data, which should not have been part of the training data, can now also fit the best matching neuron and is therefore classified. An advantage of this method is that every new input sample again trains the map.

In the named paper a multilayered network is used. Once the lower layer has reached a certain amount of energy, the higher layer is activated. The highest layer is finally classified with a kNN-Classifier. A different approach is presented by Van Laerhoven et al. [5]. In the *Spiking neural networks* the action potentials of neurons are enlarged by a time domain. Whenever a neuron is animated, it "fires". Afterwards the potential is decreasing with time. All potentials are added and animate the next upper layer when reaching a certain threshold. This addition corresponds to the time-dependent character of acceleration data. Figure 5 shows the explained method.

5 Conclusions

An overview of different acceleration sensor-based user interfaces was given. Many different applications, which are researched in current papers, were introduced in order to show the accelerometer's potential for motion based user interfaces. Across different topics of ubiquitous computing, like motion interaction in computer games, intuitive multimedia control, or authentication, acceleration data is used to measure human caused motion. Different aspects of sensor hardware including communications are also explained in order to get a full range of understanding the sensor's capabilities.

As interpreting acceleration data is a task of pattern recognition, feature extraction and some linear and non-linear classifiers are also explained. These algorithms provide the core of every recognition system. The approaches differ in complexity and the best fitting method has to be adapted to the current situation or appliance.

For motion being a highly intuitive human behaviour, motion based user interfaces will be investigated further in the future. Acceleration sensors will take an important place in the future research, because they are small and cheap. Moreover, they provide three-dimensional data and can easily be combined with other sensors. Motion based interfaces therefore will use accelerometers in upcoming appliances. An advantage of using these sensors is their sophisticated integration in current products and experiments which makes them well accepted by users and researchers. This can also be seen at the high number of scientific papers which exist on this topic. We therefore provided a comprehensive set of references for further research.

References

[1] Kunze, K., Barry, M., Heinz, E.A., Lukowicz, P., Majoe, D., Gutknecht, J.: Towards recognizing tai chi - an initial experiment using wearable sensors. (2006) 2.1

[2] Heinz, E.A., Kunze, K.S., Gruber, M., Bannach, D., Lukowicz, P.: Using wearable sensors for real-time recognition tasks in games of martial arts - an initial experiment. In: Computational Intelligence and Games, 2006 IEEE Symposium on. (2006) 98–102 2.1, 4.2

[3] Champy, A.S.: Elements of motion: 3d sensors in intuitive game design 2.1

[4] Van Laerhoven, K., Villar, N., Schmidt, A., Kortuem, G., Gellersen, H.: Using an autonomous cube for basic navigation and input. In: ICMI '03: Proceedings of the 5th international conference on Multimodal interfaces, New York, NY, USA, ACM (2003) 203–210 1(a), 2.2, 3.1, 3.3, 4.4

[5] Laerhoven, K.V., Gellersen, H.W.: Spine versus porcupine: A study in distributed wearable activity recognition. In: ISWC '04: Proceedings of the Eighth International Symposium on Wearable Computers, Washington, DC, USA, IEEE Computer Society (2004) 142–149 1(b), 3.1, 2, 3.2, 4.6, 5, 4.6

[6] Foxlin, E.: Pedestrian tracking with shoe-mounted inertial sensors. IEEE Computer Graphics and Applications **25**(6) (2005) 38–46 2.2, 2.3

[7] Block, F., Schmidt, A., Villar, N., Gellersen, H.W.: Towards a playful user interface for home entertainment systems. In: European Symposium on Ambient Intelligence (EUSAI 2004), Springer LNCS 3295, Springer (2004) pp207–217 2.2, 3

[8] Terrenghi, L., Kranz, M., Holleis, P., Schmidt, A.: A cube to learn: a tangible user interface for the design of a learning appliance. In: Personal and Ubiquitous Computing. (November 2005) 153–158 2.2

[9] Matsumoto, T., Horiguchi, D., Nakashima, S., Okude, N.: Z-agon: mobile multi-display browser cube. In: CHI '06: CHI '06 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM (2006) 351–356 2.2

[10] Lee, J., Lim, S.H., Yoo, J.W., Park, K.W., Choi, H.J., Park, K.H.: A ubiquitous fashionable computer with an i-throw device on a location-based service environment. In: AINA Workshops (2). (2007) 59–65 2.2, 3.3

[11] Okumura, F., Kubota, A., Hatori, Y., Matsuo, K., Hashimoto, M., Koike, A.: A study on biometric authentication based on arm sweep action with acceleration sensor. Proc. Int. Symp. Intelligent Signal Processing and Communications (2006) 2.4, 3.1

[12] Mayrhofer, R., Gellersen, H.: Shake well before use: Intuitive and secure pairing of mobile devices. IEEE Transactions on Mobile Computing **8**(6) (2009) 792–806 2.4, 3.1, 4.2

[13] Veltink, P.H., Bussmann, H.B.J., de Vries, W., Martens, W.L.J., Lummel, R.C.V.: Detection of static and dynamic activities using uniaxial accelerometers. IEEE Transactions on Rehabilitation Engineering **4** (1996) 375–385 3.2

[14] Mäntyjärvi, J., Kela, J., Korpipää, P., Kallio, S.: Enabling fast and effortless customisation in accelerometer based gesture interaction. In: MUM '04: Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia, New York, NY, USA, ACM (2004) 25–31 4.1, 4.5

[15] Mantyjarvi, J., Lindholm, M., Vildjiounaite, E., Makela, S.M., Ailisto, H.: Identifying users of portable devices from gait pattern with accelerometers. Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on **2** (2005) ii/973 – ii/976 4.2

[16] Randell, C., Muller, H.: Context awareness by analyzing accelerometer data. Wearable Computers, IEEE International Symposium **0** (2000) 175 4.6

[17] Baier, V., Mosenlechner, L., Kranz, M.: Gesture classification with hierarchically structured recurrent self-organizing maps. In: Proc. Fourth International Conference on Networked Sensing Systems INSS '07. (2007) 81–84 4.6

Tangible Interaction to Enhance Users' Experience

Thibaut Tiberghien
thibaut.tiberghien@mytum.de

January 7, 2010

Abstract

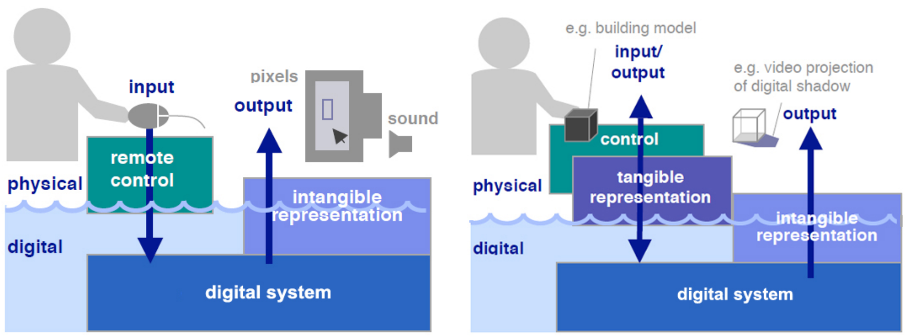
Tangible User Interfaces (TUIs) provide physical form to digital systems in order to facilitate the manipulation of bits. The main concern in the design of such interfaces is to use more efficiently the "bandwidth" of human-machine interaction taking advantage of human natural skills for interacting with his environment. There have been many research efforts put into the development of TUIs this last decade. The name 'Tangible User Interfaces' has been approved and used by a board part of the scientific community. Yet it was last decade's challenge to provide a theoretical background or organization for this field that would help the design and analysis of interfaces. This paper gives an overview of some other works written to take up this challenge. In the selection, more importance is given to papers whose focus of analysis is made on users' experience and its enhancement.

1 Introduction

While information domain was moving from the physical to the digital world, people have been surrounded by interfaces with this digital world. The most common form taken by these interfaces is a screened device like a personal computer or a cellphone. When computers first appeared, interaction with digital systems was possible only through command lines that users needed to remember and type into a console. In recent years, interaction style has largely been confined to Graphical User Interfaces (GUIs) which are a vast improvement from its predecessor — Command User Interfaces. GUIs provide a graphical representation of digital systems with pixels on a bit-mapped display using the WIMP (window, icon, menu, pointing device) metaphor. Well-known basic elements emerging from this metaphor can now be seen, pointed at and clicked on, all in order to interact with the system. GUIs have experienced a large commercial success with operating systems like Apple Macintosh or Microsoft Windows. However, people dispose of many sophisticated skills for sensing, touching, grasping, in a nutshell manipulating their physical environment that have been precluded in the context of GUIs. Limited to screens, mice and keyboards, GUIs cannot take advantage of our dexterity for numerous tasks that take place in the physical world.

Researchers have tried to develop new kinds of interfaces to counter the "intangible frustration" emerging from the limitations of GUIs. This resulted in the advent of a broad range of systems relying on physical representation of digital systems, tangible manipulation, embodied interaction, embeddedness in physical space and digitally augmented physical spaces [1]. In 1997, Ishii and Ullmer managed to unify all these systems, that may be considered as innovations proceeding on unrelated fronts, under the term 'Tangible User Interfaces' (TUIs). TUIs were defined as interfaces "that augment the real physical world by coupling digital information to everyday physical objects and environments" [2]. For them, this new kind of interfaces tends to bridge the gap between digital and physical worlds by making bits tangible. They narrowed later their definition replacing the term 'everyday physical objects' with 'artefacts', which henceforth adds more purpose to the tangibles used for interaction. They also chose to eliminate the distinction between input and output devices. TUIs then gave "physical form to digital information, employing physical artefacts both as representations and controls for computational media" [3]. TUIs have been conferred a diversity of definitions in the last decade's researches, but Ishii and Ullmer's remains the most cited and the one chosen for this work. It is also practical to consider it with its corresponding model [3][4].

The model is represented in the Fig. 1(b) and compared to the GUI model of Fig. 1(a). On the GUI model, one can note the spatial and semantic separation between input and output. The user has control of the digital system through a generic device and remotely experiences an intangible representation of it. In the context of TUIs, this



(a) A model for Graphical User Interfaces (b) A model for Tangible User Interfaces

Figure 1: Differences between GUIs and TUIs. Source: [4]

separation decreases with the apparition of a tangible representation of the digital system (or a part of it). This tangible representation is computationally coupled to the digital system and embodies at the same time mechanisms for its control. For example, in the Metadesk system [5], a small model of the Great Dome of the MIT campus (Fig. 2(a)) is coupled with this Great Dome on a graphical map of the campus. By grasping it and placing on the desk, the digital map appears on the desk's surface positioned with the Great Dome beneath its tangible representation. The user can then move or rotate the artefact to move or rotate the digital map. In this system, the digital graphical map on the surface of the desk is the intangible representation of the Fig. 1(b). This intangible representation is actively computationally coupled to the tangible representation and takes advantage of the malleability of pixels or more generally, bits. By using both tangible and intangible representations, users can use their hands to grasp a physical system in order to control more directly the underlying digital system and their experience is augmented by the display of information about the digital system that cannot be dynamically embedded in the physical system. The success of a TUI often relies on a balance and strong perceptual coupling between both representations.

2 Classifications of TUIs

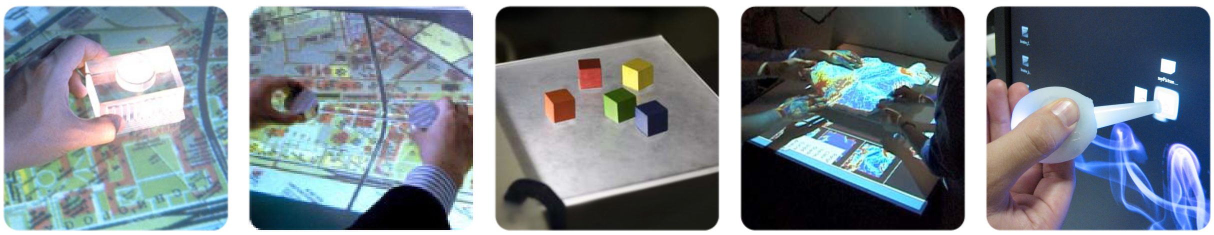
The first part of this work has introduced the field of TUIs and a model of this kind of interfaces. But as explained by Fishkin, TUIs have been largely an "I know one when I see one" field [6]. Several works have tried to build an organisation for the domain to move beyond that and will be mentioned below. It would in fact be useful to dispose of a general and theoretical framework to observe and compare different interfaces to help guide their design. Since coupling of physical and digital systems is primordial to TUIs, such a framework could help us to understand what kind of coupling is best fitted for a given application, as well as the aspects that developers should focus on in order to enhance users' experience.

2.1 Overview of selected classifications

In such a motley field as TUIs, it is not surprising that very different classifications have seen the light of the day, some based on the concept of the interface, others on the artefacts used. When Ishii and Ullmer introduced the term 'Tangible User Interfaces' [2], they already presented a division of this interaction model in three non-mutually exclusive areas — *Interactive surfaces*, *coupling of bits and atoms* and *ambient media* respectively correspond to the transformation of surfaces like walls or desktops in active interfaces, to the seamless coupling of digital systems with physical graspable artefacts and to the use of light, sound or air flow for example as representation of digital systems stimulating the human background awareness. They later provide another classification of TUIs [3] in three families based on the concept used for the interpretation of the users' manipulations of the artefacts. A typical example of the *spatial* approach is the Metadesk system [5] presented in the introduction. The *relational* approach regards adjacencies of artefacts as bindings for the underlying digital system. Finally, in the *constructive* approach, artefacts are used by the users like LEGO™bricks whose interconnections are then computationally interpreted.

Holmquist *et al.* have a more specific focus on the artefacts used for interaction and classify them into three categories [7]. *Containers* are generic artefacts used to move information between devices, a tangible version of thumb drives in a way (see for example the Slurp system describe in part 2.2). *Tokens* are artefacts that physically resemble the system they represent, like the Metadesk's dome of Fig. 2(a). And *Tools* are artefacts used to physically manipulate digital systems, like the Metadesk's scaling and rotation tool of Fig. 2(b). For this second artefact both metal disks are set on a sliding bar and coupled respectively with the MIT Great Dome and the MIT MediaLab, by moving the second disk along the bar the scaling function is activated.

The last classification that has been selected is more important in the context of this work as it is a taxonomy based on the users' impressions or feelings towards the interactive system. It is referred here to the two dimensional taxonomy presented by Fishkin [6] and cited often. The next part is dedicated to this taxonomy.



(a) Metadesk's dome representating the MIT Great Dome. Source: [5]
(b) Metadesk's constrained tool for scaling and rotation. Source: [5]
(c) Trackmate: a large scale generic tangible interface. Source: [8]
(d) Illuminating Clay: a 3D tangible interface for landscape analysis. Source: [9]
(e) Slurp: the digital "eyedropper". Source: [10]

Figure 2: Some examples of TUIs this work refers to.

2.2 Fishkin's Two Dimensional Taxonomy

The formal taxonomy described by Fishkin [6] provides a general and theoretical framework for comparing interaction systems and helping design of interfaces. It defines for this purpose a two dimensional space to represent them. This space has for axes characteristics of the interfaces called *Embodiment* and *Metaphor* that are related to users' feelings towards the system and whose signification is explained below.

Embodiment This characteristic quantifies how much the users feel that the digital system (or a part of it) is embodied "within" the physical system they are manipulating. This can also be described as the "distance" between tangible and intangible representations (see Fig. 1(b)). Four levels of this characteristic are presented:
Full – This limit case describes an interface for which no difference can be made between input and output; the physical system fully embodies the state of the digital system represented by him. The user's focus is reduced to one central and unique point. For example, the Slurp system [10] is a digital "eyedropper": as represented on the Fig. 2(e), the user can for example "suck up" a file or folder from a screen and "drop" it to another one or directly play a music file by "dropping" it to some speakers. Here the user holds in his hand the entire (digital and physical) system and for him, the data as well as the processing is embodied "within" the artefact.
Nearby – At this level, tangible and intangible representations are directly proximate to each other, and there is no need for the user to switch his focus between them. A typical example is the Metadesk [5] where the digital representation of the map appears just beneath the physical artefact when this one is positioned on the interactive surface. In this common case of TUI, the user is aware that he is not directly manipulating the digital system but the proximity is such that he feels like he was.
Environmental – In this case, the intangible representation is a so-called ambient media (sound, air, light…). There is no trouble for the user's focus; his background awareness is used here. This joins the idea of ambient media systems from Ishii and Ullmer except that for them, there is no control device which means no tangible representation. Then the intangible representation has only an informative purpose. A good example of environmental embodiment is the ToonTown system [11]. In this system, the user translates physical avatars representing people in a chat room to adjust their audio settings: a friend whose avatar is translated closer to the middle will have his voice played out louder.
Distant – For this last level of embodiment the user has to change his object of focus to switch between tangible and intangible representations. The interaction is more divided in space, manipulating the tangible representation as

control device has its result (intangible representation) displayed on another screen or even in another room. This is for example the level of embodiment of a TV remote control. In the domain of TUIs, Trackmate [8] is a good example for this case. The Trackmate system represented in Fig. 2(c) provides a generic tangible interface. Artefacts like wood cubes in the picture are tagged on their bottom face and tracked on the image of a video camera positioned under the glass surface, this can be interpreted computationally as an input device by a personal computer. Any adapted program can then be controlled in a tangible way and in many cases the result of the tangible manipulation is displayed on the computer screen.

The embodiment of the digital system within the physical system can be used to obtain a seamless interaction style as the three first levels of embodiment demonstrated. But the distant level is not to neglect, indeed Trackmate already appears like a more generic solution than other TUIs cited and a more powerful input device than a mouse or keyboard. Some advantages on these classical devices can already be cited: the multiplicity of input elements, the liberty and meaning potential of artefacts, couplings and interpretations, among others of course...

Metaphor This second characteristic classifies interfaces depending on the type of analogy made between the physical interaction with the artefacts and the digital interaction with the system they represent. This is highly related to the advantages that have just been cited, it is in a way quantifying how much the interface uses the meaning of artefacts, couplings and interpretations. It appears clearly already that a good use of metaphor will be a key for designing a seamless interaction system. Five levels have been given for this characteristic:
None – This level applies to Command User Interfaces for example where no metaphor is used in the physical interaction to represent the digital interaction. The use of mice and keyboards are indeed not reminding of the any aspect of the computational processing it controls.
Noun – A metaphor of noun occurs when the tangible system manipulated by the users resembles its associated digital system (or a part of it). Here the analogy is in the shape, the look of the physical system. To illustrate this level of metaphor, one could think of the Trackmate system [8] in a use where the tagged artefacts employed would use some meaningful couplings. For example bringing on the surface a cube holding the picture of a friend would open his personal homepage or search all emails exchanged with him. Users could also use as artefacts some small models of furnitures to position in an interior design software.
Verb – For this kind of metaphor, the tangible artefact does not look like its corresponding digital system but the manipulation of both of them is analogous. When an action is performed on the tangible, the same action is computationally performed on the digital system. Once more, one could imagine a use of Trackmate [8] where some generic artefacts like the wooden cubes would be coupled with video sequences in a video editing software. Then rearranging the order of the cubes would rearrange the order of the coupled video sequences. No analogy of noun is used but actions being performed are similar. It can also be referred here to the control of digital systems by tracking of hand movements, like the G-stalt project of MIT MediaLab (see `http://zig.media.mit.edu/Work/G-stalt`). In these projects, conventions on the interpreted gestures must be established and using a metaphor of verb can help the seamlessness and comfort of interaction.
Noun & Verb – This case is simply the union of the two previous ones. Both analogies of look and manipulation are used. This is the level of metaphor reached with the Metadesk's Great Dome model of Fig. 2(a) where the artefact resembles its digital equivalent and where moving or rotating it move or rotate his digital equivalent.
Full – At this last level there is no metaphor used in a sense that the user does not have to do any analogy effort while interacting with the interface. The physical and the digital systems are hardly differentiable to him. This is for example what has been seen with the Slurp system [10] already. This full metaphor can also be illustrated by introducing the Illuminating Clay system [9] represented on Fig. 2(d). This system is composed of a piece of malleable clay representing a landscape which is live scanned to update the coupled digital model. As users move the clay, computer-calculated characteristics of the landscape are projected on it using a color scale. To the users, no metaphor is necessary to use the system: "deforming the landscape is deforming the landscape".

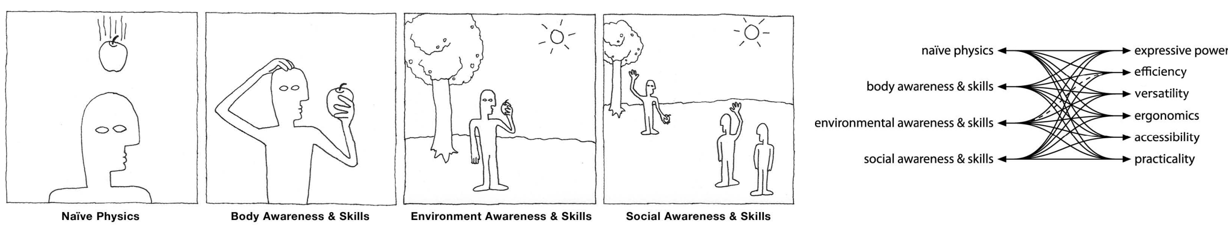
This formal taxonomy provides a useful theoretical space to represent interfaces and have an idea of users' feelings towards the system. Designers can also compare using it their work with previous ones and it helps bringing into focus the design trade-offs that have been made in the development. For example, a noun & verb metaphor makes the interaction more seamless and natural but it decreases its versatility (variety of application domains) compared to an only verb metaphor or its expressive power (variety of tasks) compared to an only noun metaphor.

3 Concerns in the design of TUIs

Classifications are useful and simple tools for researchers, for example they just need to place a given system into a two-dimensional space to draw parallels with other existing systems. However, this simplicity makes them unable to simultaneously consider many aspects of the design of new TUIs. Also, although some design trade-offs have been highlighted by Fishkin's taxonomy, those trade-offs are implicit in the taxonomy. This part focuses on other kinds of frameworks which are not exactly classifications but which also provide some useful tools for a better understanding of how to enhance users' interactive experience. These frameworks transform the disadvantages of classifications into their own advantages. They treat for example a wider part of all aspects of TUIs. They can also express design trade-offs explicitly. In few words, they use a less formal approach to make more aspects of TUIs design clearly explicit.

3.1 Overview

From the numerous frameworks aiming in a way to help enhancing users' interactive experience two papers have exemplarily been selected. First, Jacob *et al.* present a framework whose basic idea was to bring aspects from the physical world into interaction with digital systems to augment the seamlessness [12]. On their side, Hornecker and Buur provide a framework for TUIs that is focused on users' experience [1]. This second work considers a very broad field of analysis, offering several perspectives on the domain with systematic shifts of focus. This very complete work will be at the center of focus in the next part. In this part, more details will be given about the framework of Jacob *et al.* for what they call 'Reality Based Interaction' (RBI).



(a) The four 'Reality Based Interaction' (RBI) themes. (b) The six trade-offs going along.

Figure 3: Illustration of the framework described by Jacob *et al.*. Source: [12]

Jacob *et al.* organize their work around four themes emerging from the physical world (Fig. 3(a)) and they search how elements from these themes can be brought into interaction domain. By increasing the "realism" of interfaces, they tend to reduce the distance between digital and physical domains, to create a more natural interaction style and to use the numerous skills of people peculiar to the physical world. *Naïve Physics* is encompassing people's common sense of physical rules like gravity, velocity or friction. One could think of some Apple iPhone applications using metaphors of mass, gravity and inertia to improve the interface. *Body Awareness & Skills* is relating to people's abilities and familiarity with their own bodies. Making use of it offers a rich set of interaction techniques for developers. This is for example what has been used in the G-stalt project (see http://zig.media.mit.edu/Work/G-stalt). One can understand the importance of *Environment Awareness & Skills* for tangible interfaces. Indeed, referring to the "broadband" interaction of humans with their environment, this theme encompasses among others the background awareness capability used for ambient media and the ability to alter the state of physical systems by grasping, moving and so on used for "graspable" (to avoid the term 'tangible') systems like the Metadesk [5], Trackmate [8] or Illuminating Clay [9]. Finally, *Social Awareness & Skills* is related to people skills for social interaction including communication, sharing of references and objects or even collaboration. From an interactive point of view, it is an open challenge to find systems better fitted than GUIs for collaborative digital work.

This has described how connotations from the physical world can improve interaction with digital systems, yet other concerns should be taken into consideration in designing interfaces and ***trade-offs*** should be made. Six RBI design trade-offs (Fig. 3(b)) have been exposed in [12] and are summarized here. As mentioned earlier, designers might choose to use a smaller degree of metaphor and in that way make their interface less "realistic" to improve the *Versatility* and *Expressive Power* of their interactive system. Which means augment respectively the number of application domains and the number of tasks within one domain that users can perform. Users might also prefer some kind of "unrealistic" representations to gain in *Efficiency* and *Ergonomics*, in other words to be able to perform tasks (especially repetitive ones) fast and without fatigue. Finally, from a very down-to-earth perspective, designers should keep some consideration for the *Accessibility* of their interfaces to elder or disabled users for example, and for its *Practicality* including size, cost of development, durability and so on.

3.2 A Framework Focused on Users' Experience

As announced, this part focuses on the framework presented by Hornecker and Buur [1]. To provide an as-complete-as-possible analysis guide of TUIs, this framework offers four non-mutually exclusive themes that are four different perspectives on tangible interaction. Each theme has its set of concepts building possible switches of focus. If we consider again TUIs as taking more advantage of the "broadband" interaction style existing between humans and their environment, these themes and concepts are like a set of complementary "bandpass filters" that can be used to analyze one specific aspect of an interactive system. With this work, the authors are aiming for researchers to have at their disposal a more concrete guide for understanding the implications of their design choices. Themes and concepts are represented in Fig. 4 and explained below.

Tangible Manipulation This theme is of course central for TUIs and concerns the interaction with physical objects of the system. Its first concept, *Haptic Direct Manipulation*, is relating to the feeling of grabbing or moving elements in the system, which includes haptic feedback, sensory pleasure and playfulness. For example, in the Illuminating Clay system [9], it clearly brings much more comfort and pleasure of interaction that users can touch, sense and modify with their hands the surface of the landscape's physical model instead of working on a digital model of it through a graphical interface. *Lightweight Interaction* is the second concept, relating to the "conversational" style of the interaction: can users proceed in small steps and receive feedback before continuing? Looking again at the Illuminating Clay, the intangible representation that is projected on the clay surface is updated depending on its live scan, which contributes enhancing users' experience. Finally, *Isomorph Effects* concerns the understanding of the relation between a physical action and its effect on the digital system. This connects with the *metaphor* of Fishkin's taxonomy described in part 2.2. A powerful set of isomorph effects will make the interaction seamless like for Illuminating Clay.

Spatial Interaction This theme is something specific to TUIs. Although GUI style interaction can be fully realized while sitting at a desk, tangible interaction takes place in a space where users might want to move not only artefacts but also their own body to interact with the system. Here, five concepts are making up the theme. *Inhabited Space* is an invitation to think about the meaning of the place where users and artefacts meet. Interaction presents different characteristics whether it is around an augmented table where experts can collaborate, like for Illuminating Clay (Fig. 2(d)) [9], or in front of an interactive projected surface for a presentation. Also it is important to be able to appropriate the interactive space by rearranging some elements for example. This is the concern of the *Configurable Materials* concept and can be illustrated with the Trackmate system [8]. Indeed, there are a lot of different ways that one can build his Trackmate and some of those ways are described on the webpage `http://trackmate.sourceforge.net/buildit.html`. The next concept, *Non-fragmented Visibility*, refers to the ability of users to follow all visual references. For example, if someone is pointing at something, can everybody see it? Or can the users follow some gestures with their gaze without any fragmentation? This contrasts with most attempts in tele-communication. *Full Body Interaction* characterizes how much the users' skills for body expression are used. In other words, can people use their whole body as an interactive tool? And at last, *Performative Action* concerns the possibility to communicate something through the way we move while doing things. If interactive systems seldom attach much importance to these two last concepts, the G-stalt project (see `http://zig.media.mit.edu/Work/G-stalt`) is based on these concepts to establish new interactive methods and lets us understand the potential of such design considerations.

Embodied Facilitation As TUIs move computation control into the physical world, new interaction structures are defined. These structures facilitate, prohibit or limit some possibilities of actions and the behaviours going along. Of course there is a strong relation between the design, the shape of a system and its facilitations. This theme is dedicated to concepts ruling this relation. The first concept, *Embodied Constraints*, is relating to the physical system set-up in the space. Size, shape, location of objects are so-called embodied constraints that enable or disable, advantage or hinder some activities guiding users through their interaction. The Metadesk's scaling and rotating tool represented in Fig. 2(b) could have been realized by adding a model of the MediaLab to the already available Great Dome model

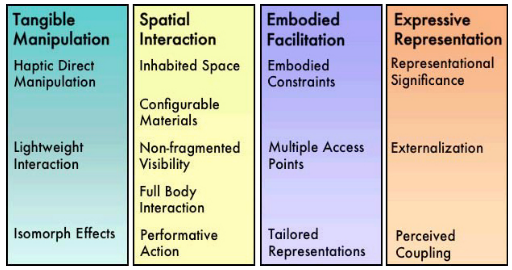


Figure 4: Themes and concepts building Hornecker and Buur's framework focused on users' experience. Source: [1]

(Fig. 2(a)). However, if developers did so, the rotation of both models on themselves would have been meaningless, they chose so to had some constraints to guide users through their manipulations [5]. Another important advantage of TUIs on GUIs is the spatially-distributed access to controls of the system. Making good use of it, an improvement of the collaboration aspect for partners using the system at the same time is possible. This is analyzed in the second concept: *Multiple Access Points*. Finally, it appears already that an important challenge in designing interfaces is to facilitate easy usage to improve seamlessness. However, it is also important that systems can gain power while building on existing skills and experience of a targeted group of users. This is the concern of the last concept: *Tailored Representation*. Referring once again to the Illuminating Clay [9] as an illustration, it can be seen that what may be judged as complex by most people turns into a powerful albeit simple tool in the hands of an architect.

Expressive Representation TUIs are based on physical representations of digital systems. Interacting with an interface means seeing and interpreting the representations of the system to react and modify its state as desired. The analysis of the representations is the concern of this last theme. Its first concept is *Representational Significance*: this is questioning the meaning power of the representations used. Whether this meaning is long-lasting in the eyes of the users must also be taken in account. For example, users might find the Metadesk's Great Dome model [5] more meaningful than a generic cube representing the same digital Great Dome in Trackmate [8] because this second representation is not so long-lasting. An important trade-off between representational significance and versatility is highlighted here. The second concept, *Externalization*, is relating to the human ability to express and share his thoughts, which makes us able to talk, teach, explain, share references and so on. Applied to TUIs, this concept refers to the help brought by the interface for discussing and coming to an agreement with partners. Can artefacts be used as props that will provide some focus to the discussion and help to make decisions? For example, Illuminating Clay [9] might be used as a prop experts could stand around for a meeting. Finally, the concept *Perceived Coupling* leads us to wonder about the clarity of the couplings between physical and digital systems. In other words, is the relation between what the user does and how the system reacts clear? This is what was called by the authors 'the faked causality' and must be strengthened to increase the seamlessness.

4 Conclusion

After analyzing the motivations for research of new interaction methods, we have characterized in this paper what means "Tangible User Interfaces". The basic idea of these interfaces is too take better advantage of the "broadband" interaction style of humans with their environment. With the objective of enhancing the users' experience with interactive systems, this paper highlights important design considerations, trade-offs and challenges with classifications and frameworks of TUIs. Besides describing the positive effects of tangible interaction on enhancing interface seamlessness, this paper also reflects on solutions provided by interactive systems for collaborative interfaces through provision of numerous artefacts. These artefacts can eventually be constrained to guide users' manipulations and present an exceptional diversity and liberty of representations, coupling and interpretations of users' actions. By leveraging on existing knowledge and skills, there are also possibilities of constructing specific interfaces. Last but not least, this paper shows how the applications of theoretical frameworks and organizations can serve in grasping better the full implications of design choices.

References

[1] Hornecker, E., Buur, J.: Getting a grip on tangible interaction: a framework on physical space and social interaction. In: CHI-CONFERENCE-. Volume 1., Citeseer (2006) 437 1, 3.1, 3.2, 4

[2] Ishii, H., Ullmer, B.: Tangible bits: towards seamless interfaces between people, bits and atoms. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM (1997) 241 1, 2.1

[3] Ullmer, B., Ishii, H.: Emerging frameworks for tangible user interfaces. IBM systems journal **39**(3) (2000) 915–931 1, 2.1

[4] Ishii, H.: Tangible bits: beyond pixels. In: Proceedings of the 2nd international conference on Tangible and embedded interaction, ACM New York, NY, USA (2008) 1, 1

[5] Ullmer, B., Ishii, H.: The metaDESK: models and prototypes for tangible user interfaces. In: Proceedings of the 10th annual ACM symposium on User interface software and technology, ACM New York, NY, USA (1997) 223–232 1, 2.1, 2(a), 2(b), 2.2, 3.1, 3.2, 3.2

[6] Fishkin, K.: A taxonomy for and analysis of tangible interfaces. Personal and Ubiquitous Computing **8**(5) (2004) 347–358 2, 2.1, 2.2

[7] Holmquist, L., Redström, J., Ljungstrand, P.: Token-based access to digital information. Lecture Notes in Computer Science (1999) 234–245 2.1

[8] Kumpf, A.: Trackmate: Large-Scale Accessibility of Tangible User Interfaces. PhD thesis, Massachusetts Institute of Technology (2009) 2(c), 2.2, 2.2, 3.1, 3.2, 3.2

[9] Piper, B., Ratti, C., Ishii, H.: Illuminating clay: a 3-D tangible interface for landscape analysis. In: Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, ACM New York, NY, USA (2002) 355–362 2(d), 2.2, 3.1, 3.2, 3.2, 3.2, 3.2

[10] Zigelbaum, J., Kumpf, A., Vazquez, A., Ishii, H.: Slurp: tangibility spatiality and an eyedropper. (2008) 2(e), 2.2, 2.2

[11] Singer, A., Hindus, D., Stifelman, L., White, S.: ToonTown: less is more in somewire audio spaces. In: Proceedings of the CHI. Volume 99. (1999) 104–111 2.2

[12] Jacob, R., Girouard, A., Hirshfield, L., Horn, M., Shaer, O., Solovey, E., Zigelbaum, J.: Reality-based interaction: a framework for post-WIMP interfaces. (2008) 3.1, 3, 3.1

Camera, AR and TUI based smart surfaces

Michael Eschey
Michael.Eschey@mytum.de

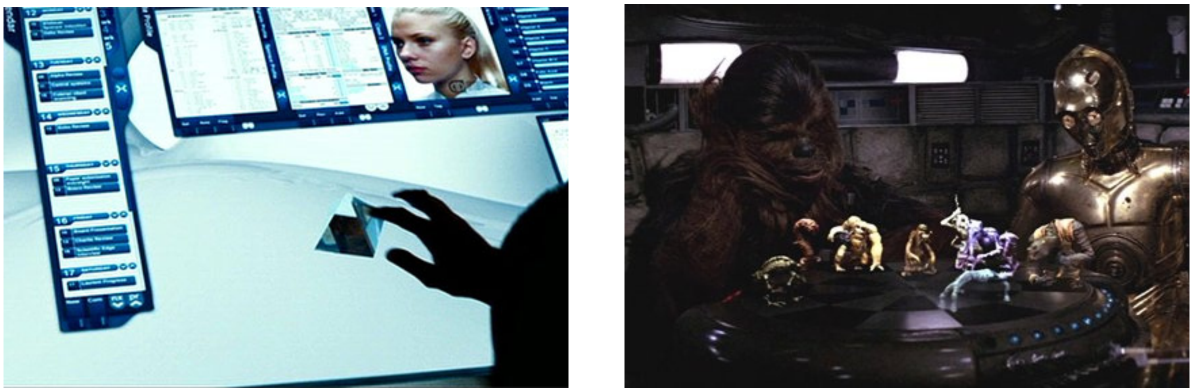
January 5, 2010

Abstract

Desktop applications, medicine, conferences, gaming industry and many other domains need intuitively usable and adaptable user interfaces. This paper gives a basic overview of smart surfaces based on cameras, tangible user interfaces and augmented reality. Therefore a selection of recent implementations will be introduced to show state-of-the-art technology and its challenges in this field of application.

1 Introduction

In the movie Minority Report actor Tom Cruise shuffles, controls and manipulates data on large displays solely with a data glove and hand gestures. Remarkably, this user interface has been developed for real by engineers [1] to be employed in the movie. It is an excellent example how small the gap between science fiction and reality is. Visions of directors and authors can be a driving factor for science and research and the humankind can profit of the advances in technology, e.g. surgery in a sterile environment. In [24] a system was developed for neurosurgeons who can browse through MRI images by hand gestures similar to the movie. Adapting to users the system features a fast learning curve for both, the users and itself. A first beta test showed, that this kind of application can be of significant use for surgeons. In [18] another medical application was presented that allows remote access to a suture on micro blood vessels via telecommunication systems. The surgeon controls two robotic arms resembling his extended hands at the surgery site using two manipulators with seven degrees of freedom, i.e. translative, rotatory and open/close. Force sensors at the tips of the slave arms and images of a moveable microscope at the site give a haptic and visual feedback to the user at the remote place. The images are shown on a moveable display which adjusts its position and zoom accordingly to changes in posture of the surgeon which are extracted from a video stream captured by a CCD camera installed on top of the display. But not only medicine can benefit from smart surfaces. Figure 1(a) shows a multi-user tabletop interface from the movie "The Island". The user controls information and data via translucent tangible user interfaces. Astonishingly, SLAPbook was introduced in [26], which is a collaborative multi-touch tabletop environment which can be controlled by translucent widgets like knobs, sliders and keyboards similar to the movie. The setup of SLAPbook is a rear projected diffuse display on which the users can place the tangibles. The basic idea of the transparent widgets is that information displayed on the table will not be occluded by them, thus offering the user more flexibility in placing and organizing his workspace. Another field of application for smart surfaces is the gaming industry. In Figure 1(b) a chess like game from the movie "Star Wars" shows the potential of augmented gaming experience. The "chessmen" on the board are visualized and brought to life by holograms. Instead of holograms, which are still difficult to realize, head mounted displays are used in [19] to augment reality for a racing game. The gaming environment is defined by a board of any size with optical markers. Additional objects can be placed in the scene by movable markers. A camera installed on the head mounted display is used for tracking visual markers and provides a video stream the tracked objects can be mixed into. The user thus gets displayed a mixed video of the scene he is viewing and dynamically added objects. Even interaction between the whole virtual car and objects tracked by markers is possible, e.g. the car hitting a windmill.
These examples show well how visions of movie producers can already be realized with today's technology and what is not feasible yet like the freely animated holograms of "Star Wars". The goal of this paper is to present techniques and state-of-the-art technology of augmented reality, tangibles and camera-based user interfaces. Hence in section 2 a selection of representative tabletop applications was made and are compared to emphasize on those intentions. In section 3 a set of techniques to realize wall-sized displays are shown and section 4 concludes this work.



(a) A multi-user tabletop diplay from the movie "The Island" controllable by tangibles, i.e. the pyramid seen on the image. Source: [2]

(b) Chewbacca and C3PO of "Star Wars" playing a holographic chess like game. Source: [3]

Figure 1: Two examples of movies using visionary human computer interfaces

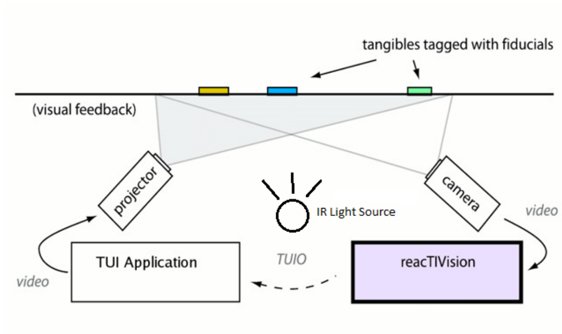
2 Tabletop Applications

2.1 TARBoard

The TARBoard in [14] is a tangible user interface and augmented reality tabletop card game environment. Conventional card games usually feature two opposing players, each having his own deck, i.e. a collection of cards. Players take turn and bring cards into play. Normally, card games do not offer boards and are only a set of rules, lacking of visual and physical feedback compared to a chess piece being moved from one field to another. TARBoard tries to enrich the players experience by visualizing the content of the cards in 3D. Thus, the gaming cards are enhanced to be tangible user interfaces. Since visual markers that are needed for the detection of tangibles can occlude information on the cards and be hidden by the hands of the players, the authors proposed a setup with a glass table. Now the markers can face down and be captured from the rear of the table. To broaden the visible region of the tracking camera, the optical path is lengthened by using a mirror. A second camera views the scene from above, but again, to get a larger viewing range under a certain angle. Hence, the whole setup needs a offline calibration. Knowing the camera parameters, objects can be added to the video stream of the top camera. Nevertheless, TARBoard reveals some inconveniences. A transparent table means that the tracking camera can view through the table and see what is behind, which can lead to false positive errors whereas the top camera can see to the ground. Additionally the stream has to be shown on a separate display. All of these factors interfere with the concept of immersion. Further, physical 3d-objects on or above the table will always be occluded by virtual objects added by the system. To overcome this problem, [13] propose two techniques for each, robust tracking of markers under partial occlusions and correct representation of depth in augmented video streams. Furthermore, as a video output device they use head mounted displays. As a reference for world coordinates, the table is furnished with a pattern of multiple fiducial markers as reference enabling interaction between tangibles instead of only relative positioning to the user. Now to firmly track tangibles despite occlusion the authors used two types of markers, rectangles and blobs. Since at least one rectangle or three blobs are necessary to determine distinct positions and accurate poses of an object, markers have to be distributed over a sufficiently large area and their relationships in world coordinates has to be well known. The problem of occluded objects of the physical world by virtual objects can be solved by calculating depth information from a stereo camera system or a depth-sensing camera [21]. Since the computational effort for those solutions is rather high, the assumption in [13] was made, that users do not expect virtual objects in scenes from the beginning and will not be disturbed if they fade in with an alpha color value. Only tangibles that are hovering above the table are expected to overlay virtual objects that affiliate with the table beneath which is easily realizable by distinguishing the markers.

2.2 reacTable

The reacTable ([10], [12]) is a multi-user tabletop musical instrument with tangibles developed at the Pompeu Fabra University of Barcelona. Figure 2(a) shows the general setup of the table. A tracking camera is placed below the table.



(a) Setup of the reacTable. Tangibles, lit by IR source (not drawn here), are placed on the table and tracked by an IR-sensitive camera, visual feedback is given by a rear projector on the diffuse table surface. [12]

(b) The reacTable in use. Tangibles with different functions are placed on the table to produce music. They can be coupled to stimulate or trigger the partner to create new effects. [10]

Figure 2: The reacTable.

To prevent the camera from viewing through the table, similar to the problem described with the TARBoard, a diffuse desk top is suggested. Tangibles will now only be detected when lying on the table and disappear for the camera directly after being lifted. If the system detects a marker, visual feedback is rear-projected. This causes the problem that tracking and projection of the visual user feedback interfere. Thus the whole backroom will be lit with infrared light which is completely invisible for the human eye. By this technique the system is able to separate between the two tasks. For the detection of IR light any CCD camera with the IR filter removed will do the work. This setup allows a compact assembly of the tabletop environment.

Beside the hardware the authors give an insight to the software part. A framework called reacTIVision was developed especially for this kind of setup offering programmers interfaces for diverse programming languages and operating systems. The framework applied on a specially designed hardware unit detects tangibles frame by frame in the video stream and sends data with the TUIO [11] protocol over UDP to client applications. For the tangibles, reacTIVision implements different fiducial engines, e.g. amoeba and D-touch [9]. First the source image has to be converted to a black and white image and then has to segmented by thresholding into regions which are scanned for the unique tree structures of the engines. Amoeba fiducials are, as the name suggests, calculated from a genetic algorithm. They form a tree structure where the nodes are represented as small circular black blobs. This enables the algorithm to detect fingers as well which are recognized on the table as a single blob. The small structure though nullifies the degree of freedom in rotation for the finger and can easily lead to false positives in the detection. Thus fingers have to be tracked over a number of frames tracing their trajectories for a robust identification. At this should be mentioned ARToolKit [4], a small framework supporting simple tracking and augmented reality functionality published under the GNU Public License.

Figure 2(b) shows the reacTable in action. Tangibles are placed on the table, each representing its own function to create sound. They can be brought into proximity to each other to influence characteristics. There are tangibles for the creation of sound, for triggering close-by function and manipulating them in different ways. Beside positioning the tangible, rotation can also be used for the manipulation of parameters, e.g. pitch of generating tangibles or frequency of triggers. This complexity requires visual feedback for the user to intuitively understand the table and work with it. Detected tangibles are visualized with auras around the physical object. Connections between tangibles are shown by waveforms or the respective form of manipulation. An additional menu can be projected around the tangibles for the use with fingers to change further parameters like the volume.

2.3 ARTable

A similar approach to the TARBoard and reacTable is suggested by the ARTable [20]. The assembly consists of a camera and projector below a diffuse table seeing the table by a mirror. A second camera with top view of the table is installed additionally for augmentation of the scene. For this kind of setup, calibration and non-occluded tracking with ARToolkit have been discussed in the previous subsections. However, the author brings up two new aspects inevitably occuring when confronted with this type of applications, i.e. keystone effects of cameras and projectors and

higher resolution tracking of moving fiducial markers. Resolution of tracking usually is restricted by the resolution of the images taken by the camera. Thus a movement predicition over time to further increase the temporal resolution is proposed. A Kalman filter with four parameters describing the 2D position, velocity and acceleration was implemented for this task. A new position is now interpolated from the new measured and estimated position.
Secondly by using mirrors or not projecting, respectively recording, orthogonally to the projection screen results into keystone effects. These effects are geometrical affine transformations, i.e. a transformation matrix can be calculated from an offline calibration step and applied as inverse to negate the influence of keystone effects in active operation. Therefore an image has to be captured by a camera which is placed orthogonal to the projection area. Now the transformation due to geometrical misalignment can be calculated from the image. The inverse of this transformation has to be applied to the video stream of the beamer beforehand to get an rectified projection.

2.4 Semi-immersive AR on Handhelds

In the introduction, a full-immersive racing game [19] was mentioned. It uses three cameras, one for tracking the controller, a second one for a top-view video stream for augmentation and a third one that is displayed for the user via a head mounted goggle. The latter is a limiting factor to these kind of applications since head mounted displays are expensive and either linked by cable or a backpack containing the computational unit, making them uncomfortable and unhandy. This restricts the amount of users who can participate in an augmented reality environment and complicates research of multi-user experiments considerably. In [25] the use of mobile phones, PDAs and tablet PCs for augmentation is examined. Cellular phones exceed concerning size, costs and spread but lack of computational power and display size compared to its overall size. Tablet PCs have the drawback of their large size and high cost. Hence, the authors in [25] concentrate on handhelds for their solution. As the core of their framework they introduce a set of existing frameworks that allow easy access to handheld APIs and offer already a great range of functionality for game developing and tracking. Since there were no 3D graphics libraries availabe for mobile devices, an own 3D-rendering framework called Klimt has been developed. The whole framework is published under an open source license in [5].
For the evaluation of their framework, the authors developed an augmented reality application for field tests, i.e. the application should be both intuitively usable for experienced and first-time users to augmented reality environments. Instead of using tangibles with fiducial markers as user interfaces, the handheld with touch screen should be used to interact with the virtual environment. Therefore a fixed tabletop has been developed with multiple fiducial markers that further permit to estimate the distance between the table and the handheld. Now a context sensitive interaction dependent on the distance is possible, e.g. giving an overview of the scene from greater distance and allowing detailed interaction when the PDA is held closer to the table. By these means, a game with virtual trains that follow a real physical, wooden track, called "The Invisible Train" was developed. The players can control switches of the track which are shown as virtual symbols on the display of the PDA and the speed of their trains. Wireless technology enables collaborative or competitive multi-user gameplay, e.g. to avoid crashes or to destroy opposing trains. To further increase immersion, trains are occluded by physical objects to give better cue of depth. The handheld itself gives an intuitive access to the virtual world implying that most people are common with the use of cellulars or handhelds. This application allows an evaluation by numerous test users and makes augmented reality accessible to broad audiences which is rarely possible for these kind of environments. It has been shown, that PDAs provide sufficient computational power for augmented reality implementations and offer a semi-immersive virtual world by looking through the handheld. This opens a completely new field of application with respect to future commercial use.

2.5 Comparison

In this section the discussed tabletop applications will be compared and evaluated. As all of the presented applications allow access for multiple users, use fiducial markers and give visual feedback, they differ in the way of implementing those key features. While keeping the tracking camera below the table seems convenient, this is not possible for handhelds. Visual feedback can be provided with back projection, additional displays, HMDs or PDAs. Table 1 gives an overview of assets and drawbacks of the examples in this section.

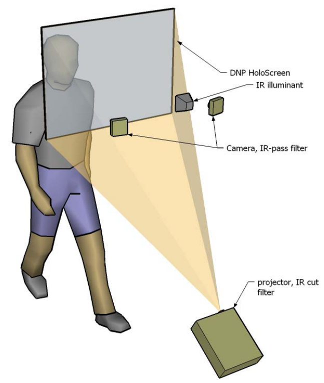
Application	Pros	Contras
TARBoard & ARTable	- no occlusion of fiducial markers - augmented reality possible with top camera - realizeable with off-the-shelf equipment - collaborative access	- either display (distracting user from table) or HMD (expensive, uncomfortable) needed for augmented stream - requires calibration - high complexity if correct depth information between real and virtual obects is essential
reacTable	- no occlusion of fiducial markers - compact assembly - visual feedback on the table itself, i.e. convenient usage - collaborative access	- requires calibration - visual feedback limited (tangibles occlude the table) - compact setup reveals drawbacks, e.g. heat generation and short optical path limiting tabletop size
PDAs	- handhelds are wide-spread - no calibration required - see-through property - PDAs offer built-in interfaces, e.g. touchscreen and buttons	- low computational power - occlusion of fiducial markers possible - limited display size - for collaborative applications each user needs his own PDA

Table 1: Comparison of the discussed tabletop applications

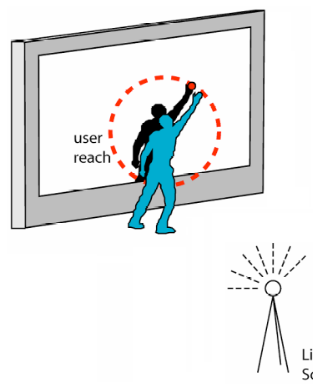
3 Gesture-based and large displays

3.1 Concept of focus and context displays

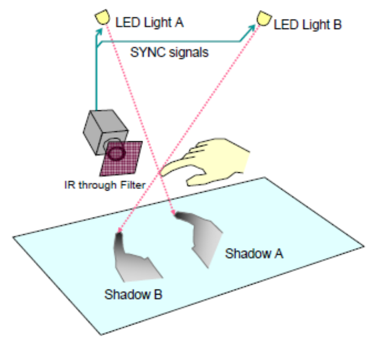
For many applications, like conferences, desktop work or brainstorming, large displays are desirable. Though no adequate off-the-shelf technology for those kind of displays is available or lack the resolution one needs to work properly. In [7] it is shown that working with multiple displays will result in splitting tasks on them with one display being the main working area. Compared to the traditional work on a desk where one normally focusses on a certain area for a task and keeps additional information accessible in the periphery, this seems natural. Secondly, users feel disturbed by spans between multiple displays resulting them to split their work on them. Another problem that arises with this application is the interaction with an array of monitors. Normal input devices like the mouse need a high sensitivity to manage the workspace meaning a lower accuracy in work and switching through applications by keyboard becomes confusing by increasing amount of data. Lastly, the user often needs higher precision for his work leading to zoom in an application where needed respectively. To avoid these problems and to adapt to the desk workspace, a focus plus context system is proposed. An arrangement of seamlessly connected multiple displays is made with one center display of high accuracy and resolution and lower resolution in the periphery reducing computational cost and attracting the user's concentration to the focussed tile. In [8] an assembly for a wall-sized display is introduced. This setup requires an own input method allowing multiple users to access the display simultaneously. As the display is divided into high and low resolution regions, the input can be as well less precise in the periphery and highly accurate in the focus area. This task is solved by a real-time finger tracking method. This is accomplished by triangulation of output streams using multiple cameras, i.e. calculating the angle between the connecting line of the camera and the finger and the optical axis of the camera. With the positions of the cameras known, the intersection of the lines can be calculated by the angles and the position of the finger can be determined. But this task is not trivial for the detection of multiple fingers. Four cameras are installed in each corner of the wall display. Each finger will now lead to one line for each camera making more intersections possible as actually fingers touch the surface. For a robust tracking two criteria are introduced by the authors. An intersection must have at least three lines without having a pair of almost parallel lines and two intersections require at least three different lines involved. As the second criterion prevents tracking in the center and a more accurate input solution in the focus area is desired anyway the authors used a smartboard [6] as input device. This makes, besides the seamless boarder between displays, a seamless handover of tracked finger data between the different input methods necessary. As the setup of the displays, three back-projecting beamers haven been chosen for the focus area and the left and right region and a beamer with a moveable projection area for the periphery. This setup allows multiple users to access the wall-sized surface. Since the tracking accuracy degrades in the outer



(a) Setup of TouchLight, projector, IR-source, two IR-sensitive stereo cameras and special coated acryl display. [27]



(b) Principle of Shadow Reach technique. Perspective projection of the shadow is used as both, tracking of gesture and visual feedback for user where and what he interacts with. [23]



(c) BrightShadow multi-user interaction technique. A camera captures shadows cast by multiple IR-sources that are time multiplexed synchronously to the cameras framerate. [22]

Figure 3: Three different gesture recognition configurations

regions circular blobs are given as visual feedback while large blobs indicate a low precision. Furthermore the interaction with objects on the wall is adapted accordingly, e.g. documents can be moved and organized in the low precision regions whereas text and content can be modified in the focus area.

3.2 Vision-based gesture recognition

In the previous section, a finger tracking system has been proposed to control a wall-sized display. It seems natural to expand the set of possible interactions by different gestures. Gesture recognition is a broad field of research concerning seamless large high resolution displays. Usual gesture tracking algorithms are based either on calculating edge images or detecting markers on fingertips, data gloves, etc. captured by one or multiple cameras. For these kinds of setups several frameworks exist for robust gestures detection, e.g. Eyepatch [16] and GART [15]. Instead of going into detail of the processing and classification of gestures which is a topic on its own [17], three vision-based gesture recognition techniques will be introduced.

[27] proposes TouchLight, a novel touch screen technology. It consists of a back-projected acryl surface with a special coating on its rear making it insusceptible to ambient light. An IR-source lights from the back of the acryl board and a stereo set of two IR-sensitive cameras installed behind the board capture streams looking through the acryl plate. Figure 3(a) demonstrates the assembly. Since the cameras see all objects behind the displays, but only touching objects and hands are of interest, the stereo setup comes into play. In an offline calibration step, the homography between camera and surface coordinates has to be determined. Now the images taken from the cameras can be rectified such that corresponding pixels of the left and right camera map to the same coordinate on the surface. For the case of wide angle lenses a distortion correction step is necessary beforehand. Now basically, to detect positions where the board is being touched the corresponding intensity values in the left and right rectified image should be the same, i.e. merging both images by multiplying results in bright areas where objects are close and dark areas for objects and persons that are distant. Further processing of the image follows traditional gesture recognition, edge images can be caluclated and classified for gestures on the board. This technology enables the construction of large, really seamless touch displays of any size.

Although, as well as finger tracking, TouchLight only allows close display interaction. In [16] the Shadow Reaching technique is introduced. The idea is basically quite simple. A bright light source is placed at a certain distance of the large display. Users now throw shadows on the display. Figure 3(b) show an user close to the display. The size of the shadow and thus the reach on the display depends on the distance relations between display, user and light bulb, meaning the greater the distance of the user to the display the larger is the shadow and reach. However a greater reach comes with the tradeoff of lower resolution and thus a lower interaction accuracy. Nevertheless, this fits quite well with the theorem of focus plus context from the previous section. From the distance organizing work can be done while

more precise interaction is possible closer to the display. Furthermore the shadow is a natural feedback of one's own actions on the display.
In [22] another shadow based gesture recognition system is suggested. It resembles the finger tracking method with four cameras, but instead of using multiple cameras, only one IR-sensitive camera is used that captures the shadows thrown by fingers from multiple IR-sources that are synchronized with the camera (see Figure 3(c)). To detect multiple fingers or hands the shadows thrown have to be assigned to the corresponding light source. To easily obtain this relationship, the lightsources are time-multiplexed on the video frames. And finally to substract ambient light from the images one reference image with all lights off is needed as well. The determination of 3D-position now follows the calculation similar to stereo camera systems. Instead of using IR, visible light can be used as well and multiplexing by using different wavelengths. The advantages of this approach are significant. Only one high-speed camera and a few LEDs are sufficient to build an user interface that allows for remote control of large displays. It can be applied orthogonally to the display viewing direction creating an invisible wall of interaction and allowing simple projection from the front as almost any application discussed until now needed rear-projection.

4 Conclusions

In the introduction some movies were mentioned showing futuristic user interfaces. But science fiction and real science are often closer than expected. In Section 2 a set of tabletop application based on tangibles and augmented reality have been presented. If one compares solely the images from "The Island" and the reacTable (Figure 1(a) and 2(b)) it is almost impossible to distinguish between fiction and reality. Some aspects might be difficult to realize yet though. The TARBoard from Section 2.1 resembles the holographic chess from Star Wars quite well, though for 3D-representations another output device is needed instead of holograms. But in some cases we can go even further as the movie suggests. While in Minority Report a data glove is needed to handle the large displays, with the BrightShadow approach of Rekimoto [22] discussed in Section 3.2 a whole seamless interaction with wall-sized displays seems possible. So in this paper it has been shown, that today's technology for human computer interaction is a fast developing field of research. For this purpose a subset of augmented reality, tangibles and camera-based approaches has been chosen to show state-off-the-art technology for this kind of application.

References

[1] http://www.oblong.com. 1

[2] http://www.istartedsomething.com/wp-content/uploads/2007/06/surfacedeskisland2.jpg. 1(a)

[3] http://www.thinkthinks.com/wf/wp-content/uploads/2009/10/starwars.jpg. 1(b)

[4] http://www.hitl.washington.edu/artoolkit/. 2.2

[5] http://studierstube.icg.tu-graz.ac.at/handheld_ar/. 2.4

[6] http://www.smarttech.com/. 3.1

[7] P. Baudisch, N. Good, and P. Stewart. Focus plus context screens: combining display technology with visualization techniques. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, page 40, 2001. 3.1

[8] Sebastian Boring, Otmar Hilliges, and Andreas Butz. A wall-sized focus plus context display. In *Proc. Fifth Annual IEEE International Conference on Pervasive Computing and Communications PerCom '07*, pages 161–170, 2007. 3.1

[9] E. Costanza. D-touch: A consumer-grade tangible interface module and musical applications. In *Proceedings of Conference on HumanComputer Interaction (HCI03)*, pages 8–12, 2003. 2.2

[10] Sergi Jordà, Martin Kaltenbrunner, Günter Geiger, and Marcos Alonso. The reactable: a tangible tabletop musical instrument and collaborative workbench. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Sketches*, page 91, New York, NY, USA, 2006. ACM. 2.2, 2(b)

[11] M. Kaltenbrunner, T. Bovermann, R. Bencina, and E. Costanza. TUIO: A protocol for table-top tangible user interfaces. In *Proc. of the The 6th Intel Workshop on Gesture in Human-Computer Interaction and Simulation*, 2005. 2.2

[12] Martin Kaltenbrunner and Ross Bencina. reactivision: a computer-vision framework for table-based tangible interaction. In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 69–74, New York, NY, USA, 2007. ACM. 2.2, 2(a)

[13] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana. Virtual object manipulation on a table-top ar environment. In *Proc. IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, pages 111–119, October 5–6, 2000. 2.1

[14] W. Lee, W. Woo, and J. Lee. Tarboard: Tangible augmented reality system for table-top game environment. In *2nd International Workshop on Pervasive Gaming Applications, PerGames*, volume 5. 2.1

[15] K. Lyons, H. Brashear, T. Westeyn, J.S. Kim, and T. Starner. Gart: The gesture and activity recognition toolkit. *Lecture Notes in Computer Science*, 4552:718, 2007. 3.2

[16] D. Maynes-Aminzade, T. Winograd, and T. Igarashi. Eyepatch: prototyping camera-based interaction through examples. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, page 42, 2007. 3.2

[17] S. Mitra and T. Acharya. Gesture recognition: A survey. *IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS PART C APPLICATIONS AND REVIEWS*, 37(3):311, 2007. 3.2

[18] M. Mitsuishi, S. Tomisaki, T. Yoshidome, H. Hashizume, and K. Fujiwara. Tele-micro-surgery system with intelligent user interface. In *Proc. IEEE International Conference on Robotics and Automation ICRA '00*, volume 2, pages 1607–1614, April 24–28, 2000. 1

[19] Ohan Oda, Levi J. Lister, Sean White, and Steven Feiner. Developing an augmented reality racing game. In *INTETAIN '08: Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive enterTAINment*, pages 1–8, 2007. 1, 2.4

[20] Y. Park and W. Woo. The ARTable: An AR-based Tangible User Interface System. *Lecture Notes in Computer Science*, 3942:1198, 2006. 2.3

[21] WA Redmond. Depth-Sensing Video Cameras for 3D Tangible Tabletop Interaction. 2.1

[22] J. Rekimoto. Brightshadow: shadow sensing with synchronous illuminations for robust gesture recognition. 2008. 3(c), 3.2, 4

[23] G. Shoemaker, A. Tang, and K.S. Booth. Shadow reaching: a new perspective on interaction for large displays. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, page 56, 2007. 3(b)

[24] JP Wachs, HI Stern, Y. Edan, M. Gillam, C. Feied, M. Smith, and J. Handler. Real-Time Hand Gesture Interface for Browsing Medical Images. *Int. J Intel. Comp. Med. Sci. Image Proc*, 1(3):175–185, 2007. 1

[25] D. Wagner, T. Pintaric, F. Ledermann, and D. Schmalstieg. Towards massively multi-user augmented reality on handheld devices. In *Third International Conference on Pervasive Computing*, pages 208–219, 2005. 2.4

[26] Malte Weiss, Julie Wagner, Roger Jennings, Yvonne Jansen, Ramsin Khoshabeh, James D. Hollan, and Jan Borchers. Slapbook: tangible widgets on multi-touch tables in groupware environments. In *TEI '09: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, pages 297–300, 2009. 1

[27] A.D. Wilson. TouchLight: an imaging touch screen and display for gesture-based interaction. In *Proceedings of the 6th international conference on Multimodal interfaces*, pages 69–76, 2004. 3(a), 3.2

Capacitive Sensing Technology in User Interfaces

Xiao Xu
Technische Universitaet Muenchen
Arcisstrasse 21, 80333 Muenchen, Deutschland
xiao.xu801@mytum.de

Abstract

Capacitive sensing is a technology of detecting displacement, force, velocity, ect. It bases on capacitive coupling effects. As a human interface device(HID) capacitive sensing is becoming more and more popular. Many products such as computer monitors, cell phones, laptop trackpads, MP3 players etc. are using this technology. In this paper we detail the principles of capacitor, capacitive sensing and how to detect the changing of capacitance in many ways. Then we present some human interface device that based on capacitive sensing technology such as Thracker - a hardware which can track hand gestures in 3D in front of a screen without touch, captable - a wood table equipped with capacitive sensing and it can track hand and body motion above and around the table. Capshelf - a shelf equipped with capacitive sensing, it can monitor where people reach and track the count of items that in the shelf. Other devices such as SmartSkin will also be mentioned.

1. Introduction

1.1. Background

In our daily life, we take a book from a shelf, we type words on the keyboard, we move a box on a table. Those are traditional way to interact with things. If we can control things or interact with computer systems just by our body, eyes or other gestures without touch, what will happen? The novel input technology will answer this question. It's a revolution of the traditional communication ways and it will greatly change our life. This technology has important applications in embedded interactive systems. Sometimes the system must know what the users want to do, e.g. turn on the book, push a button, room pictures in a touchpad and so on. By using sensors the system can recognize some of the human's gestures. In this field capacitive sensing and electric field sensing are widely used.

1.2. Current Applications

For many years people are using capacitive sensing to measure the distance between two metal plates or the overlap area of them. This technology has extremely high precision. A displacement of nanometer range can be detected. It is also used to measure the thickness of a metal foil.

Another application of capacitive sensing is to measure the fluid level. If we put an object which has a different dielectric constant from the original value between two capacitor plates, the capacitance changes. When the fluid level changes, the dielectric constant(fluid/air) between the capacitive sensing changes too. This causes a changing of the capacitance. The capacitance changing is nearly linear with the changing of fluid level.

This paper is organized as follows. Section 2 is some related work. In Section 3 we introduce the principles of capacitive sensing, including the properties, measurements of capacitance and the 3 modes of electric field sensing. New researches of capacitive sensing user interface devices are presented in Section 4. Potentials for application are described in Section 5. Section 6 concludes the paper and outlines future work.

2. Related Work

The first capacitive sensing user interface device is the Theremin. It's an instrument invented by Leon Theremin(Fig.1).

The Theremin has two antennas, one is like a circle and the other is a straight rod. By changing the distance between one hand and the circle-shape antenna a player can adjust the volume.By approaching the rod antenna with the other hand the pitch of musician will be changed. Playing the theremin is a challenge because the theremin has no physical feedback which tells the player what tone he is playing. In addition, if the environment changes,the system must be compensated by the player exactly. However, the theremin is the unique musical instrument which is played without physical contact

Figure 1. The Theremin. Source: www.wikipedia.com

Today the technology of capacitive sensing has become part of our computers, e.g. the touchpads of the laptops. However, the input is limited to a very small range. Extending the range and tracking gestures are still technical challenges.

Smith et Al. and Zimmerman et Al. developed a contactless hand tracking device using electric field sensing. The device can measure the disturbance of electric field by human hand. Although electric field sensing has higher resolution than capacitive sensing, it requires significantly more hardware and processing.

In 2001 Jun Rekimoto invented the GestureWrist. Here the gestures can be discriminated by the changing cross-section of the wrist. There are sensors put around user's wrist. When user changes his gesture,the cross-section of the wrist changes.The muscles moving beneath the skin leads to the changing of distance between sensors and skin, and it causes the changing of capacitance. Then some gestures can be discriminated.

SmartSkin (2002) takes the some idea of the touchpad, but it has one more dimension.When a user's finger approaches the grid, it shunts off some of the capacitance generated. From the amount of capacitances shunted off the distance between finger and SmartSkin can be measured. Detection range is about 10 cm. SmartSkin can track multiple fingers, but cannot distinguish users.

Capacitive sensing user interface is not a new idea. But due to some disadvantage(metioned in section 3) they are not overall in our daily life.Meanwhile in the human computer interface, especially gesture recognition, there are very few results discussed that make use of capacitive sensing.

3. Principles and Theory of Capacitive Sensing

By putting two metal plates close together without touch, it forms a simplest capacitor. Voltage difference between the two plates causes a current. The current makes the capacitor charged and it stores energy. Remove the voltage source and connect the to plates through a circuit, the stored energy initiates a current. So a capacitor is like a charge accumulator. The capacitance depends on the area and distance of the plates , and the dielectric between them(see eq.(1)). The capacitance can also be calculated with voltage difference and the stored energy(see eq.(2)).

$$C = \varepsilon_r \varepsilon_0 \frac{A}{d} \tag{1}$$

$$C = Q/V \tag{2}$$

In equation (1), C is the capacitance, ε_0 is the dielectric constant of vacuum (8.854187817F/m), and ε_r is the related dielectric value of the medium between the plates. A is the area of the plate and d is the distance between the two plates. In eqution (2) Q is the charge that the plates have and V is the voltage between the two plates. Table 1 shows some dielectric values for common materials.

Table 1. Common dielectric values. Source: www.wikipedia.com

Material	Related Dielectric Value
Vacuum	1
Air	1.00059
Polyethylene	2.25
Paper	3.5
Rubber	7
diamond	5.5-10
Silicon	11.68
Water(20 $°C$)	80.1

3.1. Definition of Capacitive Sensing

The changes of some physical quality such as displacement, force, velocity, will cause the changing of the capacitance.There are many definitions about capacitive sensing. E.g. Capacitive sensing consultant Larry Baxter mentions on his website(http://www.capsense.com):

Capacitive Sensors sense many different qualities by measuring capacitance. Variables like motion, humidity, acceleration, position, proximity, fluid level, and material composition.

The US National Highway Traffic Safety Administration (NHTSA) explains in their final rule on Federal Motor Vehicle Safety Standards:

Capacitive sensing means the detection of an object by the measurement of a disturbance in an electric field.

But there is a big difference between capacitive sensing and electric field sensing(EFS),though both of them are using the properties of capacitors and electric fields.In capacitive sensing we just measure the capacitance changing, while in EFS we need to detect the electric field disturbance by the objects or objects motion. In Raphael Wimmer's diplom thesis he gave a clear definition of capacitve sensing and electric field sensing[7]:

Electric Field Sensing (EFS) is a technique to detect an object or determine its distance to an electrode by measuring disturbance in an electric field. For distance measurement the amount of capacitive coupling between the object and one or more of the electrodes is measured.

Capacitive Sensing is a technique to measure the distance between an electrode and a conductive object by measuring the capacitance of the capacitor created by them. Capacitive sensing is a special case of EFS.

3.2. Capacitance Measurement

One of the most important things for capacitive sensing is how we can measure the capacitance as exactly as possible. There are various methods,such as passive way and active way.

For measuring the changing of capacitance in passive way, the capacitor is fully charged and a voltage is applied to it. The changing of capacitance leads to a current. To measure the displacement of current the changing of capacitance can be easily calculated. Because the current is linearly dependent on the capacitance changes.

For measuring capacitance in active way we measure the fully charge time of a capacitor in a resonant RC or RL circuit. The speed that the capacitor fully charged depends on the maximum capacitance. This is the most common way to measure a capacitance.

Here we introduce another two ways to measure a fixed capacitance.

A. Constant Current Measurement

For measuring a fix capacitance we start from equation (2). Change (2) to (3) and differentiating both sides with respect to time, results Equation (4).

$$Q = C \cdot V \tag{3}$$

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = C \cdot \frac{\mathrm{d}V}{\mathrm{d}t} + V \cdot \frac{\mathrm{d}C}{\mathrm{d}t} \tag{4}$$

As the capacitance is fixed, so $\frac{\mathrm{d}C}{\mathrm{d}t} = 0$. And in the left part of equation (4), $\frac{\mathrm{d}V}{\mathrm{d}t}$ is defined as current i, so we have equation (5).

$$i = C \cdot \frac{\mathrm{d}V}{\mathrm{d}t} \tag{5}$$

Rearranging (5) and integrating by time t yields equation (6). (5) and (6) are another most important capacitor equations.

$$V(t) = \frac{i}{C} \cdot t \tag{6}$$

Applying a constant current to the capacitor yields a linear V-t curve. Different capacitance leads to different slop(Fig. 2).

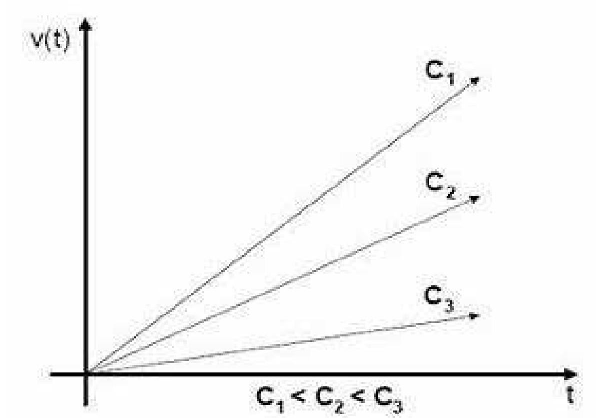


Figure 2. V-t curve on different capacitances. Source: [3]

By measuring the slop i/C we can get capacitance C. As the slop is linear, two measurements are necessary to determine the capacitor, $\frac{\Delta V}{\Delta t}$. Equation (7) shows the formulation of calculating capacitance C by knowing slop $\frac{\Delta V}{\Delta t}$.

$$C = i \cdot \frac{\Delta V}{\Delta t} \tag{7}$$

B. Relaxation Oscillator Measurement

For a more advanced solution we use a relaxation oscillator(Fig. 3).

I_{src} is a current source. It is connected to a bus and provides a constant current to charge the capacitor. One or more capacitor can be connected to the bus at the same time. The reset switch allows the capacitors to be drained. When a capacitor connects to the bus, it is charged and generates the charging voltage curve. Once the voltage reaches a predefined reference level, V_{ref}, the comparator outputs

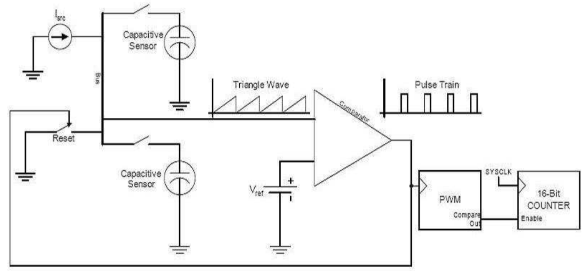


Figure 3. CapSense Relaxation Oscillator. Source: [3]

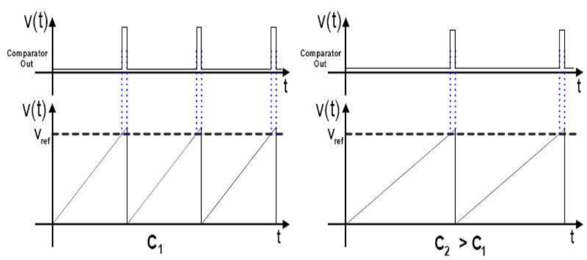


Figure 4. Charging Curve and Pulse Sequences. Source: [3]

a pulse and drains the capacitor to ground. For different capacitance we have different pulse period(Fig. 4).

The pulse period decreases with a greater capacitance. The pulse sequence is tied to the clock input of a PWM. A larger capacitance(or larger period) yields to a longer PWM output. The output of PWM enables a 16-bit counter(Fig. 5). After the PWM disables the counter, the counter value is stored and send to process.

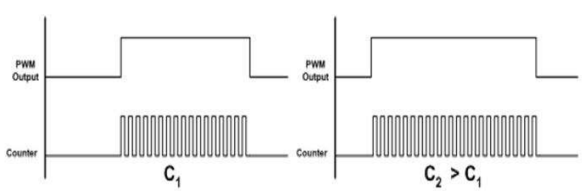


Figure 5. PWM Output and the Counter. Source: [3]

3.3. Capacitive Coupling

Capacitive coupling is an important effect of capacitors(Fig. 6). According to equation (1), a lower distance between two plates yields the higher capacitance. Capacitive coupling means, if we just apply a voltage at only one plate, the other plate through the electric field, if it connects to the ground, will also has a potential field that is similar to the first plate. The second plate receives the electric field of the first plate and form a new electric field. For objects that are not directly connected to the ground can also act as a capacitor plate if they have a good capacitive coupling to a grounded electrode.

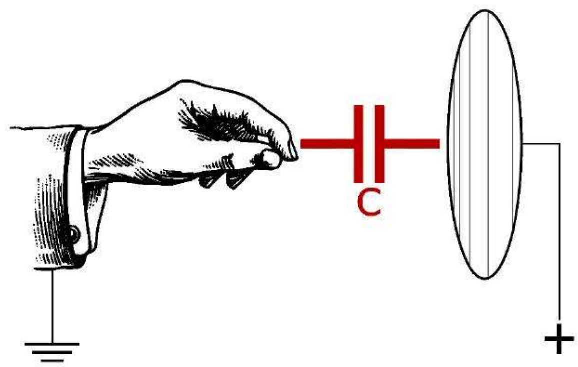


Figure 6. Capacitive Coupling. Source: [7]

3.4. Electric Field Sensing

Capacitive sensing is the first widly used technique, which utilizes the capacitance changes to measure distance. However, by studying the principle behind capacitive sensing, another sensing techniques is found and described by Joshua R.Smith. It's named Electric Field Sensing (EFS). A closing object can disturb the electric field and the EFS measures the changes of electric fields(Fig.7). Smith presents a lumped circuit model of electric field sensing in his PhD thesis(Fig. 8). There are seven capacitances involved in EFS. T and R represent transmit and receive electrodes. They are each coupled to ground by C_{T-G} and C_{R-G}. Meanwhile they are capacitively coupled to each other C_{T-R}. When the user puts his hand close to the plates, the hand and each of the plates form a capacitor, C_{B-T} and C_{B-R}. The human body itself is capacitively coupled to its surroundings C_{B-G1} and, via his feet, to ground. And R_B is the body's internal resistance. There are 3 modes in the EFS circuit[7].

Loading Mode EFS:

In this mode the receiving electrode is ignored. The circuit just uses the capacitor C_{B-T}. By measuring the coupling capacitance between transmitter T and the user's hand, a distance can be detected.

Transmit Mode EFS:

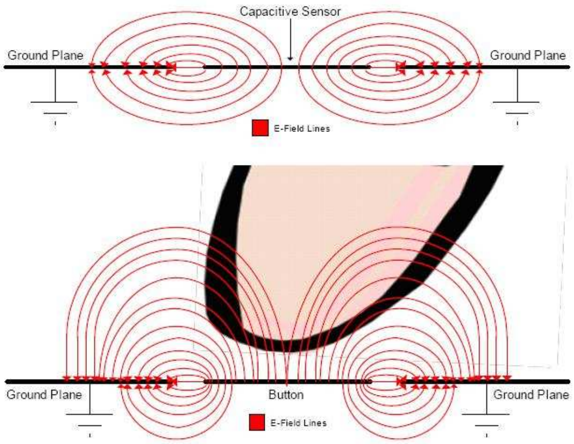


Figure 7. E-Fields and the Disturb. Source: [3]

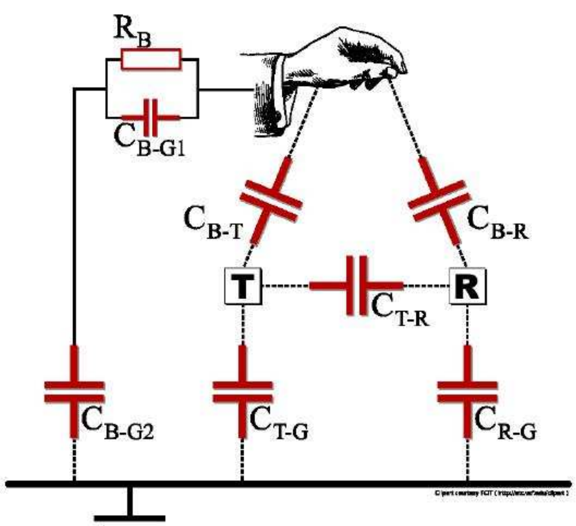


Figure 8. Lumped Circuit Model of Electric Field Sensing. Source: [7]

When EFS works on this mode, the user's hand is connect to transmitter electrode T. So C_{B-T} should be very low. By measuring the displacement current the strength of electric field can be dectected. By measuring capacitance of C_{B-R}, the distance between hand and receiver can be determinated.

Shunt Mode EFS

In shunt mode C_{T-R}, C_{B-T}, C_{B-R} are all in same order of magnitude. The electric field sensing measures the amount of electric field strength that the hand shunts off to ground. When the user's hand close to transmitter T, C_{B-T} increases, resulting C_{T-R} decrease. So the displacement current at the receiver R gets lower. To measure the current change the EFS can detect the distance between hand and Transmitter T. Because of the use of distinct T and R, the transmit mode and the shunt mode EFS have a higher resolution than loading mode EFS.

4. HID Based on Capacitive Sensing

In this Section we introduce some prototype systems, including harware and software. Captoolkit is developed by Raphael Wimmer in [5]. It is a hardware(capboard) and software platform for capacitive sensing.

4.1. Thracker

The thracker is a prototype system that allows 2D gesture input. We attach four sensor plates around a PC screen, and the sensors are connected to the Thracker device on the back of the screen. The thracker is connected to PC via USB.

A. Hardware

A NAND gate, a 300 kOhm resistor and a sensor plate roughly consists a resonant circuit(Fig. 9). The capacitance increases if the user's hand approaches the sensor plate, according to the principle of resonant circuit, the frequency of the circuit decreases. By measuring the frequency we can dectect the distance between hand and the plate of capacitive sensing. Usually depending on the distance of hand, the range of resonant frequency is at 60kHz - 120kHz. This signal is input into a 14-bit binary ripple counter. A clock signal controls whether the counter is enable or not. The count will be sent to the computer for processing by a USB-IO-chip.

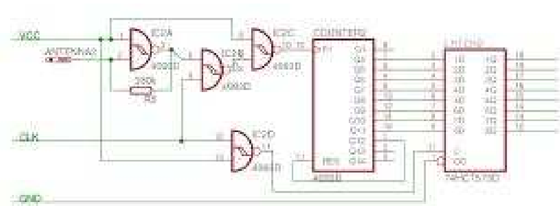


Figure 9. Hardware Circuit of Thracker. Source: [4]

B. Software

The data from USB-IO-chip are processed by a Java application using a standard API. The software calculates the distance between the hand and each of the sensing plates.

C. Expanding to 3D interaction

Although the system allows only 2D input, we can additionally define two modes which allow 3D gesture input(pick and drop). In the 3D mode we need to calculate the Z axis value using eqution (8).

$$Z = d_1 \cdot arccos(\frac{w^2+d_1^2-d_2^2}{2wd_1}) \quad (8)$$

Whereas d_1 and d_2 are the distance between the hand and sensor plates and w is the distance between two sensor plates.

However, the Z axis data is not always clear. For example, if a user pulls back his hand out of the sensors, the sensors will firstly recognize this gesture as a movement in Z axis, such as "Pick". One solution is to add a delay. The gesture causes a changing in display only the hand stay within the sensor range the whole time. The Z axis data could also be used for a mouse click. When a movement within 3cm, the sensors regards it as a mouse click, while far away from 3cm regards as a normal Z axis movement.

D. Advantages and Disadvantages of Thracker

The capacitive sensing based Thracker has many advantages.In the following we highlight some issues.

- Low cost sensors: The Thracker needs only some standard ICs and a commonly available USB interface chip. The total cost of the prototype is under 15 Euro.
- Small size, robustness and invisibility: The size of Thracker device is about a cigarette box. The sensor plates are 4 cm wide. The Thracker doesn't need moving parts, so it can be embedded into other devices.
- Scalability, high precision and speed: Additional sensors can be easily installed. The precision of capacitive sensing is in micrometer range. And because of the very little acquire data, the capacitive sensing can run very fast.
- Ease of use: Usually the users needn't additional care and training for using the Thracker.

However, there are also some limitations. The disadvantage of capacitive sensing are:

- Sensitivity quickly decreases: the resolution highly depends on the distance. By decreasing the distance the precision decreases too.
- Objects interference: other objects that need not to be tracked can disturb the resolution. One solution is to add a shield.
- Limited and ambiguous data: As the return data is only the capacitance, the capacitive sensing sometimes can not distinguish ambiguous data, e.g. one person in front of the sensor or two persons a little far away.

4.2. Captable

CapTable is a wooden table equipped with eight steel sensor plates which are affixed to it from underneath(Fig. 10). These sensor plates are connected to CapBoard, which connects to PC and offers the sensors' communication with PC over TCP/IP.

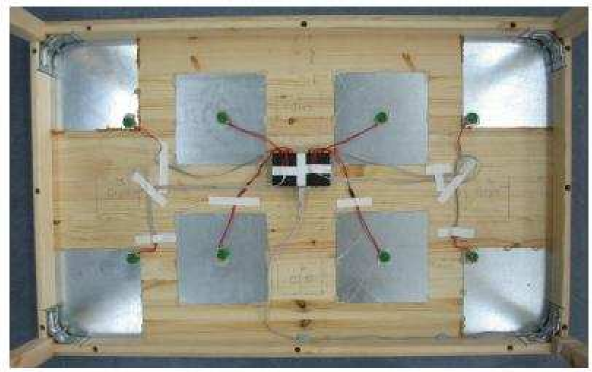


Figure 10. Bottom View of CapTable. Source: [5]

Captable can track hand gestures up to 70 cm above the table vertically and horizontally. It can also track if and where users stand around the table.Additionally, more than one hand can be tracked simultaneously, if these hands are moving with in a distinct area. In addition, CapTable's sensors can also track non-conductive objects, such as glass or porcelain. However, to recognize gestures from multiple users has not pursued yet. Because sensors cannot discriminate different user's hands.

4.3. Capshelf

CapShelf is a standard wooden shelf equipped with capacitive sensors. It can track objects inside, e.g. hands and items. In fact, the idea of CapShelf comes from the Thracker and it expands the idea. Depending on the material and distance, the capshelf can track the number of items in it. Tracking a hand with a resolution of 5cm is also very easy.

4.4. Smartskin

A. Overview of SmartSkin

The SmartSkin is a new sensor architecture based on capacitive sensing and a mesh-shaped antenna. It is an interactive surfaces that are sensitive to human hand and finger gestures. It can track multiple hand and fingers positions and detect the shape.

B. Architecture and Principles

The architecture of smartskin is shown in figure 11. The vertical and horizontal wires are transmitters electrodes and receiver electrodes. When there is a wave signal in transmitter, the receiver receives also a wave signal. Because every crossing point acts as a capacitor. If a conductive object, e.g. human's hand, approaches the crossing point, because of the capacitive coupling, it drains the wave signal, resulting the received wave signals to become weak. The position and shape of the object can be detected by measuring this effect.

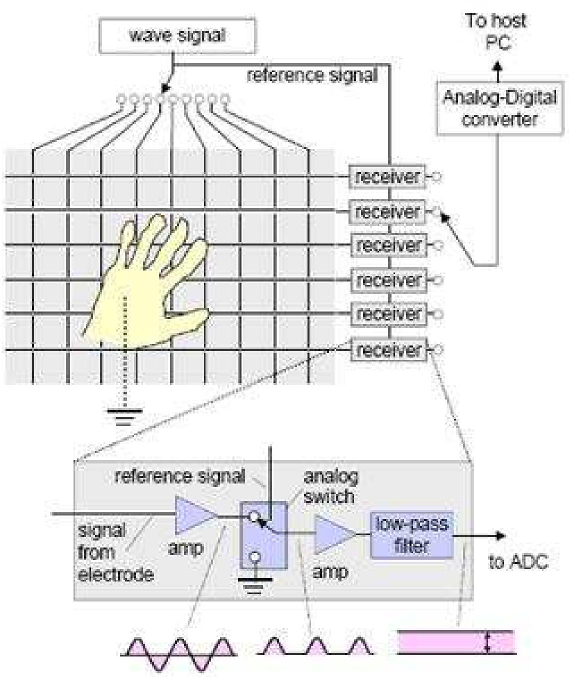


Figure 11. Architecture of SmartSkin. Source: [6]

C. Prototype 1: an interactive table

The table system is constructed by attaching sensor elements to a wooden table. A mesh-like antenna, made of polyurethanecoated 0.5 mm-thick copper wire, is laid on the tabletop.

User's hand can be detected within 5-10 cm from the table. When the hand touches the surface of table, a potential field is created. By measuring the peak of the potential field, the position of hand can be detected. There are some interaction techniques for this prototype.

- Mouse emulation with distance measurement:

  For a 2D mouse emulation, a threshold value of the distance is used to distinguish between pressed and released states. The user touches the table surface with his palm, "press" is activated, while fingers' touching means moving without mouse button pressing.

- Shape-basedmanipulation:

  In this mode a potential field is used as an input to move objects. For example, one can sweep the table surface with one's arm. Two arms can be used to "trap" and move objects(Fig. 12). By changing the hand's position around the object, the direction and speed of the object's motion can be controlled.

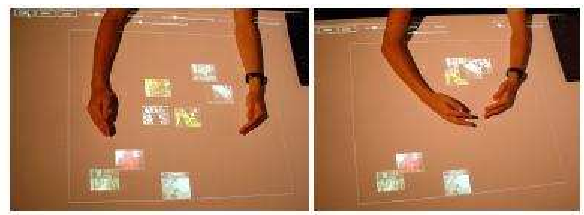


Figure 12. Shape-based object manipulation. Source:[6]

Figure 13. Map Browsing System. Source: [6]

D. Prototype 2: a gesture-recognition pad

Using a sensor with a finer grid we can determine the position and shape of objects more accurately. Additionally, we can even track multi-finger gesture. Figure 13 shows a map browsing system. The user scrolls and zooms the map with one or two fingers. If the user touches the surface with two or more fingers, by changing the distance from the fingers to the surface, he/she can control the map scale.

The gesture "pick" and "drop" can also be tracked in the prototype. Two fingers moving towards the center of an object represent "pick" and an opposite motion means "drop".

E. Discussion

The new sensing architecture can turn a wide variety of physical surfaces into interactive surfaces. However, there are also some technologies that can be developed in the future:

- Using a non-flat surface as an interaction medium: Places of interaction are not limited to a tabletop, armrests or table edges are also possible.
- Combination with tactile feedback: Currently, a SmartSkin user can receive only visual feedback, but if SmartSkin could make the surface vibrate by using a transducer or a piezo actuator, the user could "feel" as if he/she were manipulating a real object.
- Data communication between the sensor surface and other objects: The information received by smartskin can be transmitted to other devices such as a PDA or a cellular phone The table could also encode and transmit a "secret key" to mobile devices on the table, and these devices can establish a secure network with this key.

5. Applications

5.1. Thracker

While most computers now are using mouse as input devices, the thracker can provide another way for users to interact with computers. For example, in museums, exhibitions or public places, screens equipped thracker allow people to directly manipulate without other input devices. Users can interact with art, underground maps or timetables in a simple way. In addition, the traditional touch screen is difficult to use for users who have to wear a protective gloves. But the thracker enable them to interact with screens or PCs.

5.2. Captable and Capshelf

CapTable and CapShelf are perfectly suited for many different scenarios. For example, when different parts of a product are taken from different places on the table or the capshelf, the devices can track them. Thus it would be possible to optimize component placement, detect errors, or create activity patterns of the assembling person.

5.3. Smartskin

This technique can widely used on touch screen in computer or cellphone. Also it is useful in museums and public information system, such as map browsing system, mouse emulation, fingertip detection, gesture and objects' shape tracking and so on.

6. Conclusion and Future Work

In this paper we introduce some capacitive sensing based user interfaces. Those techniques are different from traditional way. We can consider them as a new input device, or input bridge to the computer. For traditional input way, we use mice, keyboard, microphone as input devices. But in this technique our hand and arm themselves are part of input devices. So we say that is a revolution of the communication technologies.

In the future we have still lots of work to do. We must continue to develop our devices in order to satisfy users' needs. Improving hardware design can improve the sensor resolution.There are several way to improve the devices. For example,

- Smaller form-factor of sensors
- Cheaper and more robust connectors
- Shielding driven by a virtual ground
- Better connection for electrodes
- Additional flash memory for off-line data logging
- Transmit mode or shunt mode support
- 3D or more dimension gesture tracking

References

[1] Capacitive sensor operation and optimization (how capacitive sensors work and how to use them effectively). *Capacitive Sensor TechNote LT03-0020*, http://www.lionprecision.com/tech-library/technotes/cap-0020-sensor-theory.html.

[2] B. Buxton. Multi-touch systems that i have known and loved. (http://www.billbuxton.com/multitouchOverview.html).

[3] J. Peterson. Capacitive sensing 101. *published in Elektronik*, October 2006.

[4] M. K. R. Wimmer, P. Holleis and A. Schmidt. Thracker - using capacitive sensing for gesture recognition. *In Proc. 26th IEEE International Conference on Distributed Computing Systems Workshops ICDCS Workshops 2006*, page 64, July 04-07 2006.

[5] S. B. R. Wimmer, M. Kranz and A. Schmidt. Captable and capshelf - unobtrusive activity recognition using networked capacitive sensors. *In Proc. Fourth International Conference on Networked Sensing Systems INSS '07*, pages 85–88, 2007.

[6] J. Rekimoto. Smartskin: An infrastructure for freehand manipulation on interactive surfaces. 2002.

[7] R. Wimmmer. Capacitive sensing for instrumented environments. *Master's thesis, Ludwig-Maximilians-Universitaet Muenchen*, 2006.

Game User Interfaces

Maximilian Mayer
maxmayer@mytum.de
Supervisor: Michael Eichhorn

January 7, 2010

Abstract

This paper describes the variety and advancement of game input devices. In consequence of the development of new technologies, augmented reality gets increasingly important in our everyday-life. Furthermore, it creates endless new and fascinating immersive gaming experiences. From traditional card- and board games, across different kinds of video game controllers, to handheld devices, funny interfaces and innovative Neurocontrollers are presented. Here the boundaries between virtuality and the real world almost vanish as a sideline, while more and more revolutionary game concepts are being developed today and in the future.

1 Introduction

Since the dawn of civilization, humanity - as well as many animal races - has been fascinated by playing games. Be it as a reason to prepare for life, hunting, to socialize or simply in order to entertain themselves during long winter nights. Along with the first desktop PCs and game consoles, digital computer games appeared on the market in the seventies. Since then, humans from all over the planet started to lose themselves for hours in the fantastic world of computer and video games. Thanks to the fast-paced development of computer hardware in the last few years, game designers are able to create highly immersive scenarios and worlds: players can visit imaginary and surreal worlds any time they want to, discover far away places and long gone ages. People are able to play all kinds of roles, and by the way test their problem solving skills on many levels.
While there have been very few different types of interfaces at the beginning, the abilities to control computers and games are now becoming as multifaceted as the games itself. This paper introduces some new ideas and interesting concepts for Game User Interfaces.

2 Game User Interfaces

Due to the multitude of devices and to get a general idea, consequently four or rather five types of interfaces are classified:

1. Implemented board- and card Games and their enhancements

2. Innovative video game controllers

3. Portable hand held interfaces

4. Funny user devices and possibly trendsetting, revolutionary concepts

2.1 Digitally Augmenting Traditional Game Environments

Whereas traditional board games always use to provide strong and direct social interaction with other people, computer games often lack this interaction and physical activity. Anyway, nowadays in the age of high-technology these differences are starting to become indistinct. New technologies and so-called smart environments connect traditional board- or card games with computers to support them or even create a whole new game-type. Thanks to pervasive computing players can enjoy the tactile feedback of game figures while they benefit from the advantages of a computer game.

2.1.1 Design guidelines for augmenting existing table top games

In their article, Steve Hinske and Marc Langheinrich [1] establish and discuss a set of guidelines for designing and implementing such augmented game environments, based on existing table top games. For them, the advantage of *augmented games* lies in the possibility to

- automatically check for rule consistencies and violations, and to
- provide players with relevant information and rules fast and in situ.

The idea is to make the gaming experience more enjoyable and convenient for the players. Technology relieves them from having to read the rules over and over again or from other time-consuming but necessary organizational duties. This way, everybody can focus on the game and the socializing aspects of it, like chatting with other players. In consequence, the authors' two main goals are not to disturb the natural game experience, plus sustaining the rich social interactions of the original game. Table 1 shows the design guidelines they use to reach these objectives.

1. The technological enhancement should have an added value.
2. The supported actions and tasks need to be clearly specified.
3. The focus should remain on the game and the interaction itself, not on the technology.
4. Technology integration should be done in a way that is unobtrusive, if not completely invisible.
5. The game should still be playable (in the "traditional" way) even if technology is switched off or not working.
6. Design and implementation should be tightly coupled.
7. The technology should be reliable, durable and safe.
8. Players should receive simple and efficient access to information. Feedback should be immediate and continuous.
9. The added technology should support the high dynamics of the game environments.
10. Development should follow an iterative process, including rapid prototyping and testing.
11. The operation of the integrated technology should be as maintenance-free as possible.
12. Secondary user interfaces should be minimized.

Table 1: Design guidelines for physical augmentation [1]

With *Warhammer 40K*, the two authors were choosing a very complex miniature war game to augment. It consists of up to a few hundred different units, tanks, buildings and a battlefield from the size that it can fill up a whole room. On top of that, several rule books make sure that the players might easily spend a few hours reading and preparing. And even during the game, they have to stop playing every now and then to measure distances or look up special rules for a particular case. One can now understand, that there is a big relief in using technology to make the game work more fluent.
Due to the use of Radio Frequency Identification (RFID) technology, the original *look-and-feel* character of the battlefield, the units or the dice stays untouched.
After an initial registration step, the system can measure distances and even orientation if the game object is

multi-tagged. This is important, because a tank can only fire in certain directions, for example. The algorithm seemed to work fine in the test scenarios, which is why almost every criteria in the list above is met. The only one that had to be broken is no. 12, due to a screen that is used to display the positions of the units and context-relevant information.

2.1.2 Smart Playing Cards

Another example for ubiquitous computing comes from Kay Römer and Svetlana Domnitcheva, ETH Zürich[2]: they equipped a classical card game with RFID chips and connected it to a desktop PC and a set of PDA's. Every single card could then be detected by an antenna that is mounted underneath the table. Besides functionality like score counting, winner determination and cheating alarm, the PDA's could be used to learn the rules and show the player whether his move is good or bad. This works even before the player makes the move. Pulling out the card a little bit is enough, and a smiley on his PDA indicates the quality of the move.
Mainly, this is just another example of the development of a software infrastructure for ubiquitous computing applications in general.

2.1.3 Weathergods: tangible interaction in a digital tabletop game

Of course, new technology also creates whole new types of hybrid-games. The digital tabletop game *Weathergods*[3] for example, combines the advantages of both traditional board games and computer games. Tangible objects can be moved on the *Entertaible touch-sensitive tabletop gaming platform* [4], which Philips developed. This digital board provides both visual and auditory feedback simultaneously. In this way, the colour of the board tiles changes and light conductors in the objects are used to transfer light into the figure of the player.

2.2 Controller

Although the earliest patent of a video game is from 1947 ("Cathode ray tube Amusement Device" [5]), it was not until the late 70's, that video games became commercially successful. At first video arcade games, then the consoles together with their controllers were designed for only one reason: to play games.
Since then, constant evolution was the crucial factor that nowadays a new era of video game controllers is about to begin.

2.2.1 VoodooIO Gaming Kit

Most of computer gaming devices can be categorized in either generic or specific. Generic devices - where keyboard and mouse rank among - can be used to play a large number of computer games, although they may not be perfect for a single one of them. Specific gaming devices, for example flight simulator cockpits, are invented to provide an ideal match to a very particular type or instance of a game, but are completely useless for any others.
With the 'VoodooIO Gaming Kit', Nicolas Villar *et al.* [6] introduce a real-time adaptable gaming controller, which is not limited in its end-user configurability. It allows players to compose and recompose their gaming space in order to fulfill their needs and personal preferences. This is made possible through flexibility of physical form and adaptability of configuration. Therefore it consists of several components, both hardware and software.

On the hardware side (as shown in Fig.1), several *pin controls* **(2)** are being put on a *substrate fabric* **(1)**, which can be easily cut to size per hand and affixed like a sticker onto any surface. The *controls* include dials, knobs, sliders, buttons, switches and a joystick that can be freely attached on the substrate fabric. Various *cables* **(3)** interconnect the different substrate pieces and a *substrate-to-USB connector* **(4)** attaches the system to a computer. Finally, the *software* allows player-defined mappings and easy interfacing with existing games. Without trouble, one can create an ergonomic gaming space that looks like the cockpit of a mech (for example to play Microsoft's *Mech Warrior 4* [7]). In this way, the VoodooIO Gaming Kit contributes to an even more immersive gaming experience where you may forget you are sitting at an office desk. Other usage they introduce in the article is a setup that reflects the design of *World of Warcraft* [8] graphical user interface with tool bars around the edges of the screen.
In a final experiment, 18 participants should evaluate the Gaming Kit compared to a keyboard with a developed,

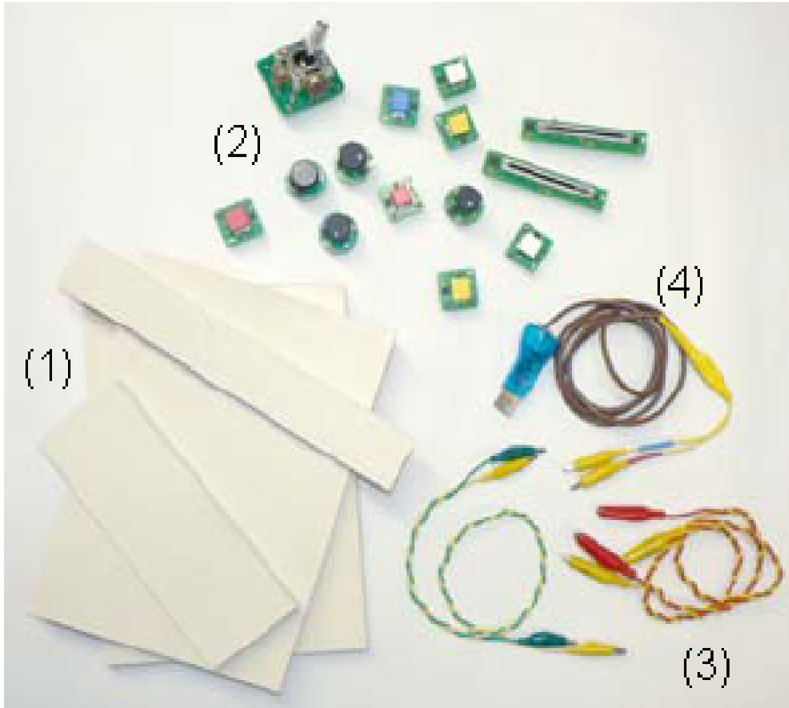


Figure 1: The hardware components: substrate fabric (1), pin controls (2), and cable interconnects (3,4). [6]

simple two-player cannon game. In the end, they could use the pin controls they wanted to and rearrange them during the game. As a result, most of the players, especially the non gamers, liked using the pin controls and the gaming pad. Although experts and some casual gamers did prefer using the keyboard to manipulate the controls, it is probably because they were more familiar with the keyboard.

2.2.2 Wii

The Nintendo WiiRemote [9] provides motion sensing capabilities integrated into a tangible remote control device. Using the WiiRemote, game players can achieve embodied interaction within the range and accuracy constraints provided by the sensing technology. This form of interaction provides a greater freedom of body movement than single-user mouse interaction in front of a computer screen. In addition, this interface offers multi-user game experiences with the help of its wireless physical interfaces. Nintendo named the system 'Wii', which sounds like 'We' and is represented with a combination of two 'i' characters, inspired from the idea of a wireless (Wi-Fi) gaming service. This trend can be interpreted as the return to physical devices, where we control things through physical body movements, by turning, moving, and manipulating the appropriate mechanical devices.

While most games are still competitive in nature, Hyun-Jean Lee *et al.*[10] were exploring prototype applications that can engage participants in active and expressive art creation in a collaborative manner. In these applications, three inter actors can work together to compose both images and sounds.

2.2.3 Sensing GamePad

Another interesting concept is the 'Sensing GamePad'[11], where measuring the electrostatic potential of a user is used to detect foot motion like stepping or jumping (which is shown in Fig. 2)

Without requiring any additional sensor-orientated input devices, these game pads are easy to set up and not sensitive to external noise, like differences in lighting conditions. So they combine the custom of the use of a normal game pad while being able to recognize human body motion in addition to game pad manipulations. Furthermore, this technique can be applied to virtually any types of input devices that contact the human body. Portable Music Players,

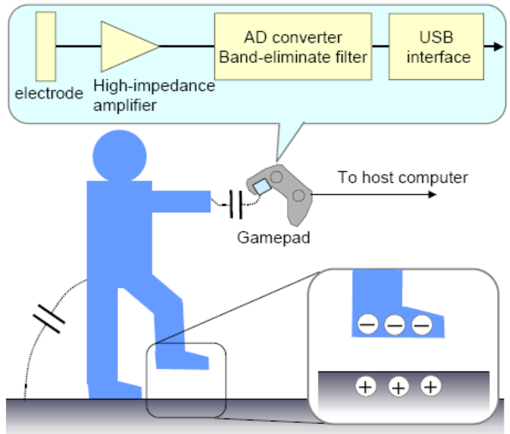


Figure 2: Principle of the proposed sensing architecture. The electrostatic charge caused by human body motion is sensed by the input device held by the user. [12]

Musical Instruments or Sports Equipment are just some ideas and it can as well help rehabilitating people with physical disabilities.

2.3 Portable hand held interfaces

Since it is common for the majority of the population in the developed countries to carry small computers in form of mobile phones with them, of course there is a huge market of applications and games coming up that use the growing functions and performance of these hand held devices.

2.3.1 Marker-Based Embodied Interaction for Handheld Augmented Reality Games

In [13], three types of handhold augmented reality games are described and presented, that can be played with every off-the-shelf mobile phone with a camera:

1. A simple penalty kick shooting-game for product packages
2. One strategy game for tabletop boards that implements a more abstract mapping of physical orientation to virtual orientation in the game
3. And a memory game for physical cards where the phone is held upside down. Most of the time, it acts as a passive observer of the users' actions and only comes into play in specific situations.

For the author, this concept is extremely promising because of the minimal requirements in infrastructure and setup. Some software and a visual code system in the form of markers is everything that is needed to play.

2.3.2 Coping with uncertainty in a location-based game

With today's small, portable hand held devices, revolutionary new location-based game concepts arise. In *Can You See Me Now*[14], three professional performers (see Fig. 3) - equipped with hand held computers, wireless network connections, General Packet Radio Service (GPRS) and Global Positioning System (GPS) receivers - run through real city streets. 15 online players are being chased through a highly abstract 3D virtual model of the host city. Everyone can see the position of each other and exchange messages. While they have to try to stay away from the runners further than five meters, it is not allowed to enter buildings or exit the game zone. Of course, in bigger cities with high buildings and narrow streets the GPS is not always absolutely accurate, and how to deal with this effect is the main concern of the article. Two strategies are suggested:

- Hide the uncertainty so that participants are less aware of it and feel minimally disrupted by its worst effects
- Reveal the uncertainty so that participants can work with it

These ideas are already being carried forward into a second game called *Uncle Roy All Around You* [15], where audiences can take part online as well as street players (in place of the professional performers).

Figure 3: A runner with game equipment built into his jacket. [14]

2.4 Other revolutionary user interfaces and fun stuff

In that paragraph, some more interesting ideas and funny concepts are being presented. Beyond that, a futuristic neurocontroller that was released on the market just recently completes this work.

2.4.1 Fitness Computer Game with a Bodily User Interface

Based on map information, Sari Mokka *et al.*[16] developed a bodily user interface with an exercise bicycle and a virtual environment. The user can steer with buttons on the handle-bar and the surface profile of the environment influences the pedaling effort. In this way, it is possible to cycle through a 3D virtual Finnish Lappland, which makes the exercise session more motivating and rich in experiences. Also, test results showed that the steering is intuitive and easy.

2.4.2 You're In Control: A Urinary User Interface

Thanks to piezoelectric sensors in the back and a screen mounted above a urinal, users can now play interactive games while going for a pee. The authors believe that this added "interactivity to urination has valuable applications to recreation, sanitation, and education."[17] Only women will need some extra device to be standing up during urination and thus be a bit handicapped.

2.4.3 Neurocontroller - Brain Computer Interface Technology

Recently, firms try to angle for the world's attention by implementing a new interface technology called *Neurocontroller*. Thanks to one or several electrodes attached to the head, neural activity can be measured in the brain. With this technique, also known as non-invasive electroencephalography, the user is able to give commands with thoughts only. The latest model in this category was released on December 22nd, 2009 by *Emotiv*[18]. It is called *EPOC* and along with a battery and wi-fi (to connect with a computer), it uses 14 saline sensors to detect brainwaves. Furthermore, a two-axis gyroscope make it able to measure four categories of inputs (see Fig.4):

1. Conscious thoughts (Cognitive suit):
   With an initial training, six rotations and six directions plus the command ’disappear’ are teached in and can be accessed later in the game. This function is yet mostly supposed to serve as a hotkey to support the user in addition to regular controllers.

2. Emotions (Affectiv suite):
   The device tries to distinguish between the emotions ’Engagement/Boredom’, ’Meditation’, ’Excitement’ and ’Frustration’. This way, a game could adapt its difficulty and characteristics individually on the basis of the emotion of the player; And thus make the gaming experience more enjoyable and less frustrating for both beginners and experts.

3. Facial expressions (Expressive suite):
   Signals from muscles around the eye - including eyebrow and eyelid - are recognized and a character in the game can for instance mimic the player.

4. Head rotation:
   Enabled by the gyroscope, this function can above all be interesting in future technologies, especially with a 3-D-Gaming Environment.

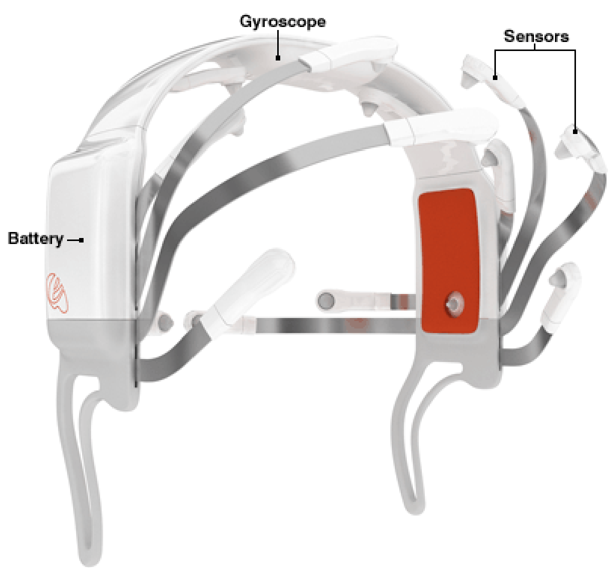


Figure 4: Sensors respond to the electrical impulses behind different thoughts; enabling a user’s brain to influence game play directly. A gyroscope enables a cursor or camera to be controlled by head movements [19].

3 Conclusions

We have shown that computer games are just as fast-paced as anything in our modern, technological world. In the last fifty years, from their appearance until today, they have gained in popularity all over the world. And not only that - they found their way into our everyday-life and it is almost impossible to get around them today. The reason for that are new ways to embed game user interfaces into our everyday-environment. Mobile phones become computers, and computers become smaller and smaller so that they can be carried everywhere, all the time. Like we have seen before, this is how ordinary cereal packages can already become part of a video game.
I am well aware of the fact that there are many more important user interfaces, like the iPhone, containing interesting technologies to mention. Therefore, the focus lies in bringing some of the more unknown technologies closer to the reader instead of just repeating what most of us already know. Anyway, nowadays it is becoming hard to sustain an overview of the available devices, hence this subject-matter going far beyond the scope of this work.
And this is only a start. In the future, board games are connected and supported respectively by computers. Then

board games itself become a controlling device. Other options include using GPS and GPRS respectively to play a game. Sometimes nothing more than the human body is needed to play a video game. In the near future even more unbelievable, striking new ideas and possibilities will come up, for example being able to control computers with our thoughts only. For many people this might already seem like magic. But we are only at the very beginning of this trend and the potential is seemingly endless.

References

[1] Hinske, S., Langheinrich, M.: W41k: digitally augmenting traditional game environments. In Villar, N., Izadi, S., Fraser, M., Benford, S., eds.: Tangible and Embedded Interaction, ACM (2009) 99–106 2.1.1, 1

[2] Römer, K., Domnitcheva, S.: Smart playing cards: A ubiquitous computing game. Personal Ubiquitous Comput. **6**(5-6) (2002) 371–377 2.1.2

[3] Bakker, S., Vorstenbosch, D., van den Hoven, E., Hollemans, G., Bergman, T.: Weathergods: tangible interaction in a digital tabletop game. In: TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction, New York, NY, USA, ACM (2007) 151–152 2.1.3

[4] Jakob Leitner, Christina Köffel, M.H.: Bridging the gap between real and virtual objects for tabletop games. In: The International Journal of Virtual Reality. (2006) 1–5 2.1.3

[5] Thomas T. Goldsmith Jr., Cedar Cove, E.R.M.: Cathode ray tube amusement device. In: United States Patent Office. (1947) 2.2

[6] VILLAR, N., GILLEADE, K.M., RAMDUNYELLIS, D., GELLERSEN, H.: The voodooio gaming kit: a real-time adaptable gaming controller. Comput. Entertain. **5**(3) (2007) 7 2.2.1, 1

[7] http://www.microsoft.com/games/mechwarrior4 2.2.1

[8] http://www.wow europe.com 2.2.1

[9] http://wii.com 2.2.2

[10] Lee, H.J., Kim, H., Gupta, G., Mazalek, A.: Wiiarts: creating collaborative art experience with wiiremote interaction. In: TEI '08: Proceedings of the 2nd international conference on Tangible and embedded interaction, New York, NY, USA, ACM (2008) 33–36 2.2.2

[11] Rekimoto, J., Wang, H.: Sensing gamepad: electrostatic potential sensing for enhancing entertainment oriented interactions. In: CHI '04: CHI '04 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM (2004) 1457–1460 2.2.3

[12] Rekimoto, J., Wang, H.: Sensing gamepad: electrostatic potential sensing for enhancing entertainment oriented interactions. http://www.sonycsl.co.jp/person/rekimoto/papers/chi04-late.pdf (2004) 2

[13] Rohs, M.: Marker-Based Embodied Interaction for Handheld Augmented Reality Games. Journal of Virtual Reality and Broadcasting **4**(5) (March 2007) `urn:nbn:de:0009-6-7939`, ISSN 1860-2037. 2.3.1

[14] Benford, S., Anastasi, R., Flintham, M., Greenhalgh, C., Tandavanitj, N., Adams, M., Row-Farr, J.: Coping with uncertainty in a location-based game. IEEE Mobile and Pervasive Computing **2**(3) (2003) 34–41 2.3.2, 3

[15] Theory, B.: Uncle roy all around you. (2003) 2.3.2

[16] Mokka, S., Väätänen, A., Heinilä, J., Välkkynen, P.: Fitness computer game with a bodily user interface. In: ICEC '03: Proceedings of the second international conference on Entertainment computing, Pittsburgh, PA, USA, Carnegie Mellon University (2003) 1–3 2.4.1

[17] Maynes-Aminzade, D., Raffle, H.: You're in control: a urinary user interface. In: CHI '03: CHI '03 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM (2003) 986–987 2.4.2

[18] Systems, E.: Emotiv - brain computer interface technology. http://www.emotiv.com (2009) 2.4.3

[19] http://news.bbc.co.uk/2/hi/technology/7254078.stmgraphic 4

User Interfaces for Media Manipulation and Creative Use

Qi Shao
thomassq@mytum.de

January 7, 2010

Abstract

Recent developments in user interfaces bring not only some new technologies with media manipulation, but also a lot of new concepts of usage to us. According to the functionalities of these user interfaces, they can be used in different fields of our daily life.

In this paper four application fields and their related user interfaces like I/O brush, VoodooIO, PaperPoint, media remote control cube and "keep in touch" are presented. With a video camera and touch sensors embedded inside, the I/O brush enables us to draw with the static and dynamic "ink" picked up from the environment. VoodooIO is a set of hardware and software components that makes the design of personal preferred control environment possible. PaperPoint and the cube serve as media remote control and make our media manipulation more convenient. At last, the "keep in touch" is a networked fabric touch screen that enables couples to feel each other over a long distance. Besides, most of these user interfaces also motivate the user to find their further potential creative use.

Keywords: User interface, Application fields, I/O brush, Malleable control structure, VoodooIO, Cube, PaperPoint, Keep in touch

1 Introduction

In recent years, there have been many efforts towards designing of user interfaces with multimedia technologies. Thanks to these user interfaces, the physical world and the virtual world are successfully connected and our daily life is becoming more and more digital. According to their functionalities, they find their way into all different application fields.

First of all, their use in the field of entertainment for children is presented. As mentioned in the papers [1, 2], with a special paintbrush called I/O brush, users can pick up all the textures, colors or even dynamic patterns from our physical environment as ink and draw on a digital canvas. This kind of user interfaces is very suitable for the children, especially for the young children who do not read and write. During the drawing, children not only understand the basic concept of elements and principles of design, but also develop their ability to sort, think and communicate. Through this visual art, it motivates people to be more creative.

In the field of our daily work, sometimes users wish a more flexible office environment because of a rigid shape, fixed layout and predetermined selection of controls. To solve this problem, a malleable control structure - VoodooIO (VIO) is introduced here. With a set of hardware and software components, VoodooIO enables users to design, deploy and adapt control structures, that follow their personal ergonomic preferences and control requirements, rapidly and dynamically [3]. This kind of user interface not only lets our work environment become more flexible, but also can make our daily works more efficient.

Media remote control is not a new topic of modern science any more. But considering convenience and ease of use, there are still some user interfaces with new ideas keeping ahead in this field. Function overload of modern systems makes button based remote controls a rather confusing user interface and some controls like TV remote controls have only a few buttons that are used frequently, while others are not used at all [5]. Here is a cube based media remote control that achieves all the necessary functionalities by different hand gestures introduced. Besides, a paper based presentation and interactive paper prototyping tool - PaperPoint will also be presented. This paper is used as an interactive medium that links to digital information and services. So the slides can be annotated in a very natural way by drawing and writing on the paper with a digital pen [6]. Furthermore, there is no need to stand near the mouse during a presentation, because PaperPoint also serves as a media remote control.

At last, there is an overview of a tactile-vision intimate interface called "keep in touch", that enable couples in a long distance relationship to feel and express intimate emotions to each other through touch, gesture and body language. Using a sensorial interface created by combining the visual and tactile senses, each partner is presented with a blurred digital projection of their lover. When they touch the body of their partner, the image comes into focus revealing their features [7].

In order to introduce all the user interfaces above, the text is separated into 4 application fields in this paper. For each application field there is a deep insight into the principle of relevant user interfaces. And lastly, some further application fields of user interfaces will be discussed.

2 Application Fields

Every user interface has its own special functionalities and creative uses. In this point, this section is separated into following 4 application fields: Entertainment for children and educational aspects, work environment, media remote control and human relationship.

2.1 Entertainment for children and educational aspects

Nowadays, parents always invest a lot of entertainment tools in children for educational purpose. A common choice is letting children begin to draw with some drawing tools or pens. But if parents expect children to learn more while drawing, there is another possibility, a new generation drawing tool - I/O brush.

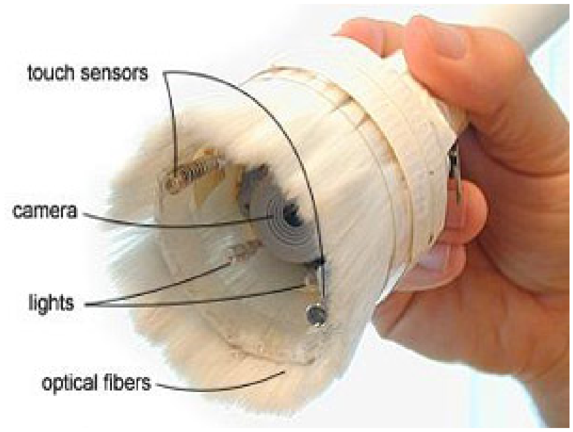
2.1.1 Functionalities of I/O Brush

Most drawing tools or pens we use have only a one-way flow of ink and in most cases, they only have control over the outflow of the ink. This means every time before drawing, all the desired colors should be created by mixing other colors and collected. It costs not only time and space, but also some colors from our physical environment are not always so easy to be mixed out.

Compared with normal drawing tools or pens, I/O brush allows both inflow and outflow of the ink. There are two modes for picking up static ink with I/O brush: color mode and texture mode. When this brush touches a surface, the color or texture of the surface will be automatically saved as ink for the next drawing step (see Figure 1 (a)). In the color mode, the system computes the RGB value of all pixels in the captured frame and returns a solid RGB value-based color. Compared to the color mode, texture mode just captures a snapshot of the surface, which consists of one frame. The advantages of these modes are that all the colors and textures from the environment can be picked up immediately. Especially for the texture, it's just a kind of ink and there is no need to draw it with different colors again and again. Furthermore, most users only need to collect some objects, whose color they need for the painting, instead of mixing out all the colors and putting a lot of palettes around them.

One another enjoyable functionality of I/O brush is its ability to create dynamic expressions that extend not only into spatial but also into temporal dimension. This mode grabs a sequence of consecutive frames of the brushed surface and then lets people draw with the captured movement [2]. There are mainly two ways to create "animated ink". The most common way is capturing an already dynamic object such as a blinking eye or a butterfly with flapping wings. Another way is applying some relative motions to the brush while capturing a static object. For example a user can create a dancing apple by twisting the brush during the process of capturing a static apple. With the combination of dynamic capturing and dynamic painting motion, the user can even create some more complex animations just like a text scroll created by brushing over the canvas using an equivalent speed but opposite direction with the ink captured by brushing over a static text in dynamic ink mode.

(a) "Picking up" different attributes of an object [1].



(b) The I/O Brush tips [1].

Figure 1: I/O Brush.

2.1.2 Structure of I/O Brush

The I/O brush system consists of two parts: the paintbrush and the drawing canvas. The brush is equipped with a small (3x3cm) CCD video camera that enables the brush to capture colors, textures and movements. Around the camera, there are several supplement light bulbs and some spring-based flexible touch sensor as shown in Figure 1 (b). If the touch sensors detect that the brush touches a surface, the bulbs around the camera will be turned on to provide enough light for the camera. During this process, the brush captures the frames from the camera and stores them in the system as ink. Besides, there are about 150 optical fibers woven into the brush. Once capturing ink is finished, these fibers will light up to give users an indication that ink is ready. On the neck of the brush there is a dial equipped, which is used for switching modes between color, texture and dynamic expression. Finally, the I/O brush should not only look like a real brush, but also feel like one. Thus, the soft acrylic hair from real brushes is bounded around the brush tip to give the user a more real feeling while drawing.

Currently, a large Wacom Cintiq screen [8] with a built-in graphics tablet is used as the drawing canvas. The coil of the Wacom pen tip is embedded inside the I/O Brush's tip to allow the system to detect the presence or absence of the brush on the canvas [1].

2.1.3 Children's work with I/O Brush

After several I/O brush studies [1] some researchers find out that some activities through the entertainment with I/O brush are very important to children, especially to ones that can not read and write yet. At the beginning, children were always curious about new colors from their environment and searched everywhere for the special color in unusual places. In this step, children not only develop their basic knowledge about different colors and their understanding of elements from their environment, but also the ability to sort and classify. After getting the ink, children drew their fantasy or a real event on the canvas very passionately. This let the children always think critically and helped them to reflect on their thoughts through abstract representations.

At last, the children were told how to create dynamic drawing. They seemed to be more excited and began with searching non-static targets like gasping mouths and blinking eyes of their partner. And then they talked to each other and shared their ideas in order to draw another picture. At the end, some of them could draw alternately with static and dynamic ink. However, combining some motion techniques for creating dynamic expression seemed still too advanced for most kindergarteners. To combine the statments above, during this time children not only develop their creativity and abstract thinking, but also the ability to communicate.

Compared with some classic education methods of parents, entertainment with I/O brush seems to be more successful. Some parents make great efforts to teach their children the connections between new concept and real life by talking about them. For children, trying to understand the new information only makes sense when the information is based on the reality that children are familiar with. Therefore, entertainment with I/O brush connected with familiar surroundings and objects can help children to understand the new concept better.

2.2 Work environment

With the development and progress of the times, computer has become the most used tool in our daily work. But its input devices have always a rigid shape, fixed layout and predetermined selection of controls, which more or less influence the flexibility of our work. To improve the work environment, here are a new kind of malleable control structure and a special tool called VoodooIO introduced.

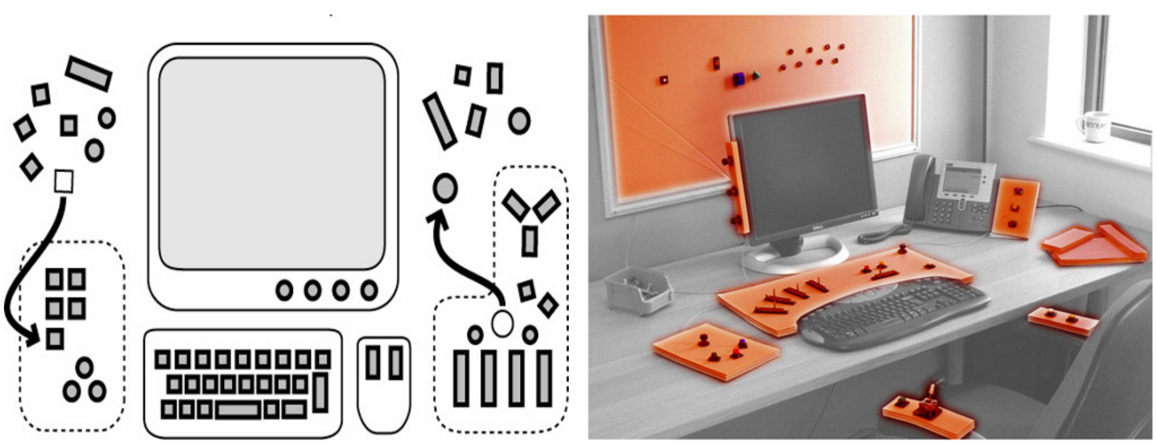
2.2.1 Malleable control structure

According to these problems above, one best way is finding some supplementary components to let them work with or even replace the computer input devices. There is a new malleable control structure that meets all these requirements. This structure is made up of a set of controls, which can be freely arranged on control areas and physically adaptable by the user. For example, control areas and controls can be introduced, organized and removed to suit interaction requirements and personal preferences [3] (See Figure 2 (a)). During this physical attachment, the components are supplied with power to operate and a network medium for communication while connected. Adding and removing the control instances are also quickly detected and registered. With all these designed elements, a user can easily make out his own work environment in accordance with his personal preference.

2.2.2 VoodooIO

VoodooIO is an implementation of a malleable control structure that meets all the requirements above. Basically, it consists of a set of software and hardware components that enables users to design, deploy and adapt control structures according to their control requirements and personal preferences.

As shown in Figure 2 (b), a VoodooIO set with special developed material is used for adding some new control areas in an office environment. Using this certain material, that covers a sector of the wall and desktop with any shape and is also affixed to the side of a monitor bezel or a chair armrest in this example, a user can easily transform the surface of furniture and equipment into control areas, on which controls can be freely arranged [3]. This material is called control substrate and works as the underlying layer on which the controls can be replaced. In order to connect the controls with each other and also some applications or systems, control substrate not only plays the role as a physical support, but also supplies enough power and digital network medium for communication or detection and propagation of interaction events like adding or removing controls. The control set of VoodooIO consists of some



(a) The input device is malleable [3]. (b) Office environment with VoodooIO design [3].

Figure 2: Malleable control structure.

basic control devices like buttons, sliders, dials and joysticks, that can be deployed ad hoc on the control substrate area. With some small pins at their base, the controls are allowed to be fastened to substrate material and connect the controls to a pair of conductive layers embedded inside the control substrate to get power and network medium supply.

Using these components of VoodooIO, users can easily change their work environment and the work is becoming more flexible. Besides, some commands to the system also require a certain duration for moving mouse and typing keyboard. But VoodooIO enables users to execute the command just through a quick touch of the controls and let the work become more efficient.

Recently, VoodooIO is also often combined with Adobe Flash [9] for design purposes. This extends the existing practice of authoring interactive applications in terms of arranging components on a virtual stage, and provides a physical stage on which controls can be arranged, linked to software components and combined with other physical design materials [4]. As shown in Figure 3, the graphical stage of the map navigation interface is always synchronized with a physical stage that supports rapid and dynamic arrangement and manipulation of the VoodooIO controls.

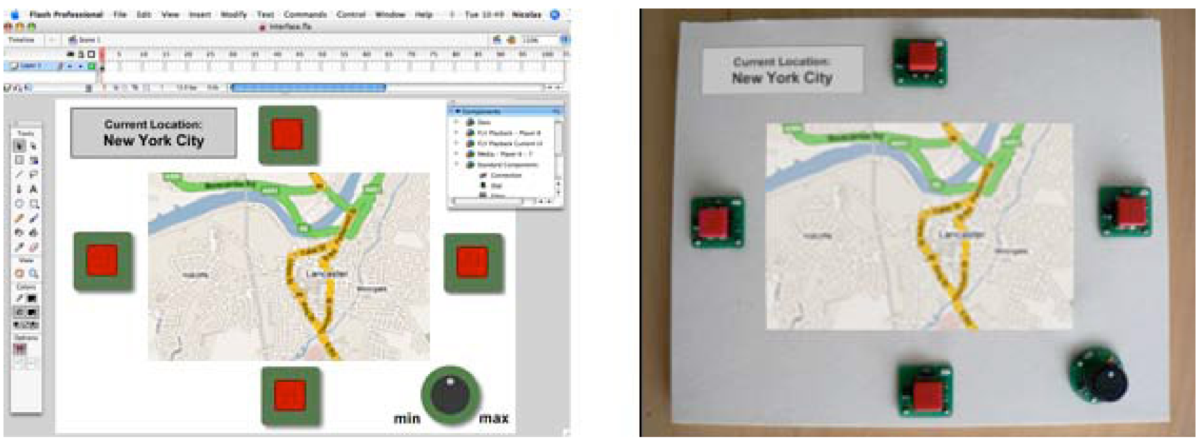


Figure 3: Map navigation interface with VoodooIO and Flash [4].

2.3 Media Remote Controls

With latest technologies, some user interfaces can also serve as media remote controls. Here are two typical examples: a cube based TV remote control and PaperPoint.

2.3.1 Cube based TV remote control

Normal remote controls like a button-based TV remote control are usually designed according to a button per function paradigm. With this paradigm, the function overload of modern systems always causes some inevitable confusion while using the remote control. Besides, there are usually only a few buttons with frequently used functions such as hopping channels and changing volume, while the others are not used at all or occasionally. Some recent television platforms even have a hierarchical menu structure that requires more concentration and operations. Furthermore, after some studies [5] researchers find out, that during operating with traditional TV remote controls some natural gestures of the human hand like grasping, holding or turning the control up to a face-up position could already be used to invoke the command itself [5].

After analyzing these factors above, a fully functional cube-based remote control is presented. According to our usual gestures of grasping and playing with some objects, the shape of the remote control is decided to be a cube with a sash positioned at the golden section of the cube as orientation support and a additional push-button on the top for distinguishing the top and bottom. The most important parts of this cube are the accelerometers and a gyroscope embedded inside, which detect the rotation and determine which side is facing up in order to read the gestures.

As shown in Figure 4, with this cube, all the most frequently demanded commands can be achieved by a set of gestures. If the button on the top is pressed, a user can rotate the cube to the left or to the right to change the volume and flip the cube up and down to switch the channel. Otherwise these gestures are also used in the set up mode to navigate through the menu. The cube can be set to the standby mode by putting it on the pedestal and shaking the cube left-right navigates to the menu home. Compared to the traditional button-based remote controls, this cube can be used without looking at the buttons and more easily with our natural gestures.

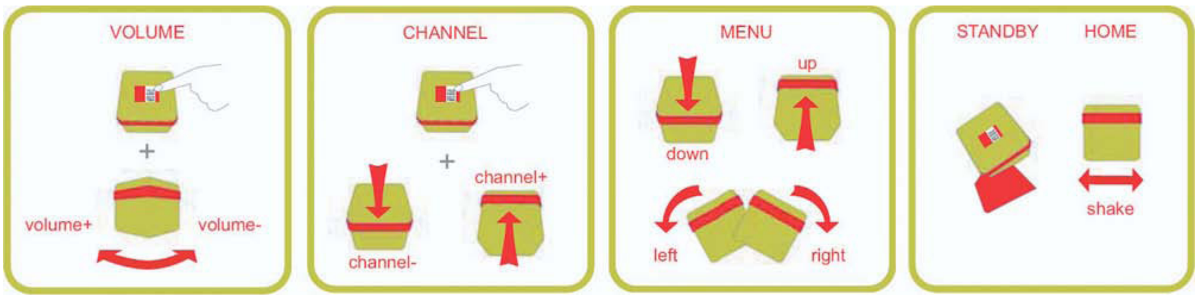


Figure 4: The operating alphabet of the cube [5].

2.3.2 PaperPoint

In some presentations with e.g. Microsoft PowerPoint, people usually stand next to the computer because they need to control the slides using input devices like mouse and keyboard. Especially when some highlights have to be pointed out, there is always a certain time they need to near the screen and then come back again to switch the slides. Besides, it is also not so easy to add some notes quickly and to do some changes like deleting and changing the position of subsequent slides in urgent cases during the presentation. Here is another media remote control, a paper-based presentation and interactive paper prototyping tool called PaperPoint based on digital pen and paper functionality offered by the Swedish company Anoto [10], introduced.

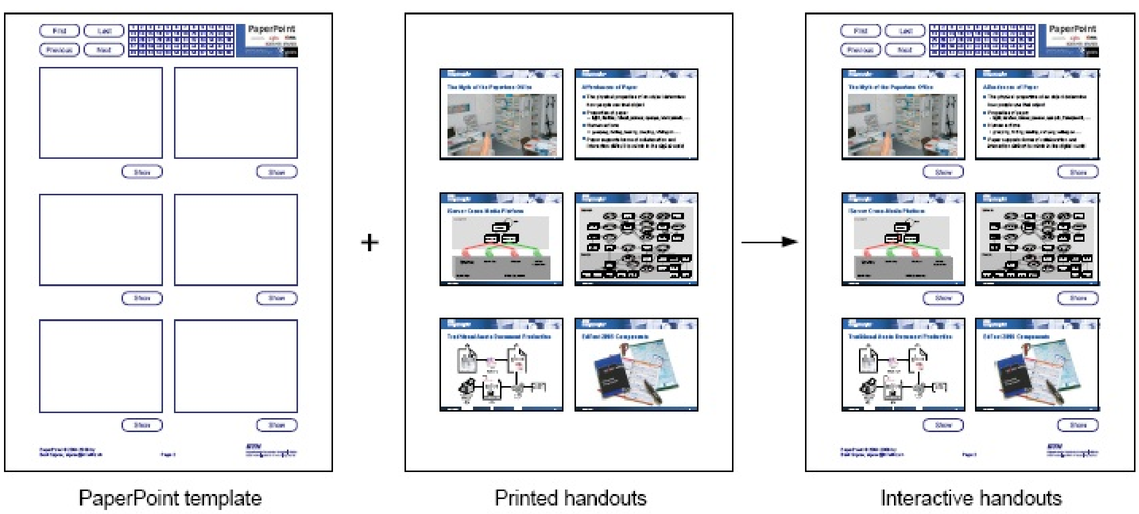


Figure 5: PaperPoint printing process [6].

Using a digital pen, the handwriting can be captured by a special interactive paper framework, which enables active areas to be defined on paper documents. These active areas can be linked to supplementary digital information or services [6]. When a user selects a position with the digital pen in these active paper region, the corresponding link is activated. Some data like (x,y) coordinates are detected and sent to the server by the digital pen. Therefore, not only the handwriting can be captured in this way, but also some commands can be executed by a simple touch on the paper within a certain area.

To create a PaperPoint presentation, first of all the slides with PowerPoint are prepared as usual. When it comes to printing the handouts, special preprocessed PaperPoint template pages have to be loaded into the printer tray in the correct order. And then the user needs to print a unique Anoto pattern together with button shapes on each PaperPoint template page. These buttons can be used to control the presentation with a light touch of a digital pen. The handouts are printed on the PaperPoint template page resulting in the interactive paper handouts (See Figure 5). After that the user needs to specify the location of the PowerPoint file and lets it combine the interactive paper handouts. From now on, if the PaperPoint client is started, the interactive paper client will wait for the input from the digital pen and the user can start the PowerPoint presentation.

PaperPoint has a lot of advantages. Holding these interactive handouts, each digital slide can be accessed via a Bluetooth connection and displayed on the screen while walking all over the place during the presentation. Another advantage is that a user can also delete or change the position of the slides during the presentation and nobody will notice that. Besides, paper is cheap, light, can easily be carried and it is always easier and more comfortable to write on paper than with some digital devices like tablet PCs.

2.4 Human relationship

With the latest media technology in the field of user interfaces, there is a new way for couples over long distance to keep their intimate relationship. This interface is called "Keep in Touch" and made up of a networked fabric touch screen that combines the visual and tactile senses together. Each couple has two of such touch screens in their own locations. Through the screen they can see the projected shadowy and blurry image of the other site. When they touch their partner's body on the screen, the image comes into focus revealing their features (See Figure 6). So the emotion of the couples can be expressed through touch, gestures and body language and sent to each other over thousands of miles.

As mentioned in [7], at each site, the translucent fabric is stretched across a doorframe with magnetic and proximity sensors embedded inside the frame. Once the fabric is touched, its edges will be pushed close to the sensors, which detect pressure and touched locations. After receiving the signal from sensors, a motor attached to a camera-projector system will be activated and brings the image into focus. When the couples depress the fabric or move their hand off their partner's body, the motor will change the focus again to show a blurred image. With this visual effect, as long as the couple touches each other, they will feel like their partner is always by their side.

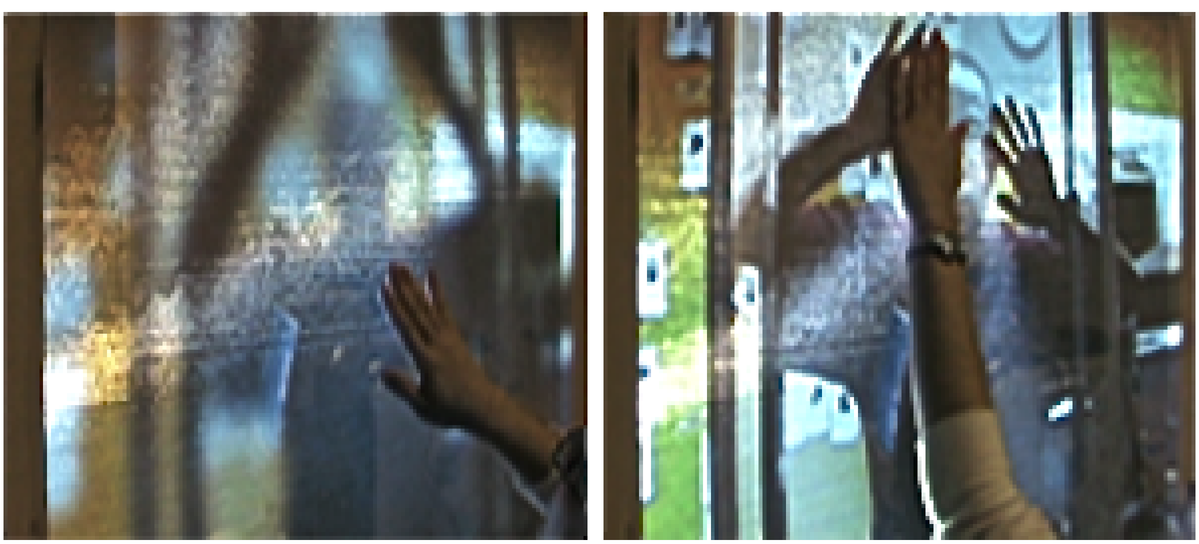


Figure 6: Keep in touch [7].

3 Further Application Fields

The user interfaces discussed above are not only used within these four application fields, but also have their further potential application fields. For example, VoodooIO can be used for designing some manipulation area near the screen or monitor at some exhibitions or in some museums. Using the components of VoodooIO, visitors can easily give some commands to the system like selecting a menu or changing the zoom. After the exhibition, the components can be taken away and prepared for the next exhibition. Another example is that the I/O brush is not just an entertainment tool for children and can also be used in the field of arts. With special aesthetic insight and skilled painting ability, creating a new art with I/O brush for the artist is just a question of time. Nowadays, user interfaces are such a sophisticated research field, that covers almost all fields of our daily life. If there is a user interface, there should also be a relevant application field and possibility of some other potential fields, that we can combine it with.

4 Conclusions

In this paper, five instances of user interfaces (I/O brush, VoodooIO, Cube based TV remote control, PaperPoint and "Keep in touch") and their application fields according to the different functionalities are discussed. With the new concept of media manipulation, all these user interfaces find their own creative use and bring many benefits to life. For example, children can learn a lot from the entertainment with that. While the work environment is getting more flexible and work is becoming more efficient, some user interfaces provide convenient remote controls. And even the problem of long distance relationships is also partially solved. Therefore, as discussed above the user interfaces cover almost all the fields of our life and have still a lot of potential fields for media manipulation and creative use.

References

[1] Ryokai, K., Marti, S., Ishii, H.: I/o brush: drawing with everyday objects as ink. In: CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM (2004) 303–310 1, 1(a), 1(b), 2.1.2, 2.1.3

[2] Ryokai, K., Marti, S., Ishii, H.: I/o brush: beyond static collages. In: CHI '07: CHI '07 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM (2007) 1995–2000 1, 2.1.1

[3] Villar, N., Gellersen, H.: A malleable control structure for softwired user interfaces. In: TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction, New York, NY, USA, ACM (2007) 49–56 1, 2.2.1, 2.2.2, 2(a), 2(b)

[4] Spiessl, W., Villar, N., Gellersen, H., Schmidt, A.: Voodooflash: authoring across physical and digital form. In: TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction, New York, NY, USA, ACM (2007) 97–100 2.2.2, 3

[5] Ferscha, A., Vogl, S., Emsenhuber, B., Wally, B.: Physical shortcuts for media remote controls. In: INTETAIN '08: Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive enterTAINment, ICST, Brussels, Belgium, Belgium, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2007) 1–8 1, 2.3.1, 4

[6] Signer, B., Norrie, M.: PaperPoint: a paper-based presentation and interactive paper prototyping tool. In: Proceedings of the 1st international Conference on Tangible and Embedded interaction, ACM (2007) 64 1, 5, 2.3.2

[7] Motamedi, N.: Keep in touch: a tactile-vision intimate interface. In: Proceedings of the 1st international conference on Tangible and embedded interaction, ACM (2007) 22 1, 2.4, 6

[8] Wacom: Cintiq. http://www.wacom.com/cintiq 2.1.2

[9] Adobe: Flash. http://www.adobe.com/uk/products/flash 2.2.2

[10] Anoto: Ab. http://www.anoto.com 2.3.2