

EnhancedTable: An Augmented Table System for Supporting Face-to-Face Meeting in Ubiquitous Environment

Hideki Koike¹, Shinichiro Nagashima¹, Yasuto Nakanishi², and Yoichi Sato³

¹ Graduate School of Information Systems, University of Electro-Communications,
1-5-1, Chofugaoka, Chofu, Tokyo 182-8585, Japan

koike@acm.org, naga@vogue.is.uec.ac.jp

² Faculty of Technology, Tokyo University of Agriculture and Technology,
2-24-16, Nakamachi, Koganei, Tokyo 184-8588, Japan

yasuto@cc.tuat.ac.jp

³ Institute of Industrial Science, University of Tokyo,
4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan

ysato@iis.u-tokyo.ac.jp

Abstract. This paper describes our design and implementation of an augmented table system for face-to-face meetings. The system was designed to be used by multiple users in the ubiquitous environment, where people do not need to bring their laptop PCs. With effective use of the advantage of computer vision, we implemented the concept of *ubiquitous desktop* as personal workspace and *virtual Chinese table* as shared workspace. User can share/personalize files by drag-and-dropping icons to/from the shared workspace. The system also provides capabilities for interactive image capturing and finger position sharing, both of which would be useful for the meetings.

1 Introduction

As personal computers become more and more and popular, changes have occurred in the style of small group meetings. Traditionally, paper documents and white boards have been the main elements in such small meetings. On the other hand, people often bring their laptop PCs to the meeting in order to take notes and to show presentation slides to other participants. In addition, people also download electronic files, such as PDF documents, from the Internet during the meeting. Other participants may copy these files from his/her PC by using file transfer tools such as FTP.

Although such digitally enhanced meetings have succeeded in introducing efficiency, we often feel a little inconvenienced. For example, laptop PCs brought by the participants require a relatively large area of the space assigned to each participant on the meeting table. Little space remains in which to place other objects such as paper documents, and it is sometimes hard for participants to write memos in their notebooks.

When exchanging electronic files, people need to know the IP address of the computer from which the files are downloaded. They have to make a connection to the computer, and copy the files onto their own PCs.

On the other hand, much research is being conducted to realize the ubiquitous computing environment proposed by Mark Weiser[20]. In the ubiquitous environment, a large number of computers are embedded in the environment and will support users' activity implicitly and/or explicitly by enabling communication with each other. If such environments were to be established, users would not need to carry their laptop PCs, which are heavy but contain important data. The users' data would be stored in network file servers which are in a safe place, and would be retrieved when necessary. It is, moreover, expected that the environment would recognize the user's position and show appropriate information at the appropriate position. However, the traditional interaction framework such as GUI, which uses a keyboard and mouse as standard input devices will not be appropriate in such a ubiquitous environment. Then the Perceptual User Interface (PUI) [17] which utilizes several recognition technologies such as image recognition and speech recognition is focused on as the next generation interaction framework.



Fig. 1. Typical scene of a current small meeting. Each attendant brings his/her own laptop PC. As a result, most of the table surface is occupied by those laptops

We previously developed a vision-based augmented desk system named *EnhancedDesk* [5]. In this project, we developed a method for real-time finger recognition [12] and a method for interactive object registration and recognition based on color information [9]. Our system, however, was designed to be used by a single user like other augmented desk systems [1, 6, 8, 21]. The vision-based interaction such as used in our system could be easily extended to multi-user environment.

This paper describes our design for an augmented table system in a ubiquitous and augmented environment and its implementation in the near future. The system aims to support effective face-to-face meeting by allowing users natural

and intuitive operation of digital files. The next section describes the key features of our design. Section 3 shows interactions in our system. Section 4 describes implementation details. Section 5 discusses advantages and disadvantages of our prototype. Section 6 concludes the paper.

2 Meeting in Ubiquitous and Augmented Environment

To date, many augmented systems have been developed. Most of them were designed for being used by a single user or for sharing one screen by multiple users. On the other hand, our interest is in integrating each user's personal workspace and a shared workspace on one screen. The key features of our system are as follows.

2.1 Integrating Ubiquitous Personal Workspaces and Shared Workspace on the Table

In the ubiquitous environment that we imagine, the user do not need to carry heavy laptop PCs which contain important files. When a user puts an object (e.g., a mobile phone) which identifies him/her on the table, the system identifies the user and displays his/her desktop with his/her personal files, which are stored in a networked file server, at that user's position. The user directly manipulates the displayed objects by using his/her own hand and fingers instead of using a mouse or trackpad.

At the center of the table, a special shared workspace is projected. When a user drags his/her personal file from his/her workspace to the shared workspace, the file becomes a shared file. Other users can copy this shared file just by drag-and-dropping the file from the shared workspace onto their own personal workspaces. Moreover, this shared workspace serves as a presentation screen when the user previews particular file as we describe later.

2.2 Enhancing Paper Documents Using Vision-Based Augmented Reality

As we described in [5], paper documents and digital documents will co-exist for the time being. The paper documents are easier to carry, easier to read, easier to add notes to, and so on. However, it is much more difficult to make copies and send them to other people quickly.

By effective use of the advantages of vision-based augmented reality, we implemented the following features which enhance the paper documents. One is interactive image capturing. The user can make a digitized image of a part of or all of a paper document just by showing a rectangle gesture. Another is finger position sharing. When the users enable the finger position sharing mode, each user's finger position is projected on the document.

3 EnhancedTable

3.1 Overview

In order to explore the interaction framework described above, we developed an augmented table system, named EnhancedTable, which allows parallel interactions by multiple users (Fig.3). The system uses a normal white table as a screen and interaction space. On top of the table, there are two LCD projectors (PLUS V-1080) for displaying images, two CCD cameras (SONY EVI-D100) for finger/hand tracking, and one CCD camera (SONY EVI-D30) for image capturing. A LCD projector, a CCD camera (EVI-D100), and a PC (Pentium 4 2.80MHz: 512Mb memory: Linux) with an image processing board (HITACHI IP5005) make a unit. Currently we have installed two units. Each unit covers a half of the table. The reason why we use two units is to get enough resolution for displaying and capturing image.

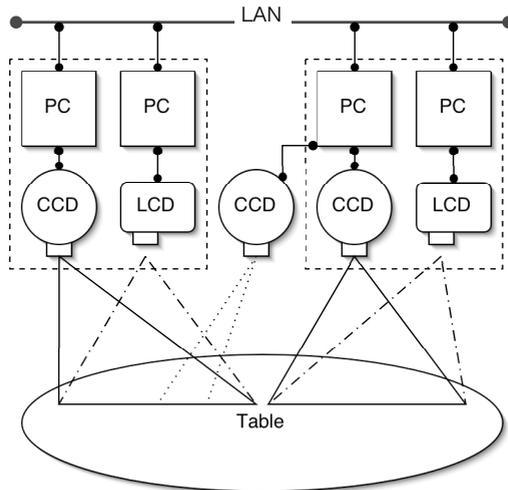


Fig. 2. System architecture of EnhancedTable

One of the important features of our system (and this feature is essential to support meeting) is the ability to allow parallel interaction by multiple users. Traditional touch panel systems do not allow such multiple interaction. Although some recent touch panel systems can detect multiple pointing, they still do not allow users to put some objects such as mug cups on the panel. It is also difficult or expensive to make a much larger table. On the contrary, our system naturally allows multiple interaction and could be extended to accommodate larger displays.



Fig. 3. EnhancedTable in use



Fig. 4. A private workspace. The mobile phone, the top-right position in this example, defines a reference point. The three icons on the left are operational icons. The four icons on the right are data icons

3.2 Personal Workspace: Ubiquitous Desktop

When the user sits at the table and put his/her mobile phone on the table, the system recognizes the mobile phone and identifies the user. Then it automatically projects the user's personal workspace such that the mobile phone is on the top-left (or top-right) corner of the workspace. Unlike other meeting systems which have embedded displays, the users can sit anywhere they want to because the projection coordinate is decided based on the position of the mobile phone.

There are data icons and operational icons in the personal workspace (Fig. 4). The data icons represent text files, image files, or audio files owned by the user. Users can drag these icons by using their fingers. The operational icons

include trash icon, preview icon, etc. The users can perform a certain operation by drag-and-dropping the data icon on the operational icon.

3.3 User Identification

To recognize a mobile phone, we used a method for object registration and recognition using an RGB color histogram[9]. We first capture the image of an object by 60x60 pixels. The registration system scans each pixel and calculates RGB values of the pixel in 256 colors. If the value of the pixel is (r, g, b) , where $0 \leq r, g, b \leq 8$, the system increments the value $H(r, g, b)$ by one. After scanning all the pixels in the image, this three dimensional matrix H is registered as a model of the object (Fig. 5). In the recognition process, the system looks for an object which has the size of usual mobile phones. When the system find the object, it calculates the RGB histogram of the object by using the same algorithm described above, and compares this histogram to those registered in the system. If the system finds a model whose histogram is close enough to the object's histogram, the object is regarded as matched to the model.

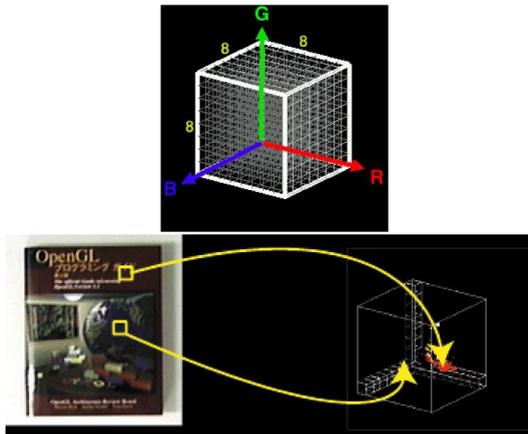


Fig. 5. RGB histogram (top) and an algorithm to make the RGB histogram. The system scans each pixel and calculates RGB values of the pixel in 256 colors. If the value of the pixel is (r, g, b) , where $0 \leq r, g, b \leq 8$, the system increments the value $H(r, g, b)$ by one. After scanning all the pixels in the image, this three dimensional matrix H is registered as a model of the object

3.4 Shared Workspace: Virtual Chinese Table

On the center of the table, a circular shared workspace is always displayed (Fig. 6). When a user drags his/her personal file from his/her workspace and drops onto the shared workspace, the file becomes a shared file. Other users can copy this shared file just by drag-and-dropping the file from the shared workspace to their own personal workspace.

The user can rotate the circular shared workspace by hand as with a Chinese dining table in a Chinese restaurant (or with a lazy susan in a regular restaurant). The user in any position can reach every file in the shared workspace by rotating that space.

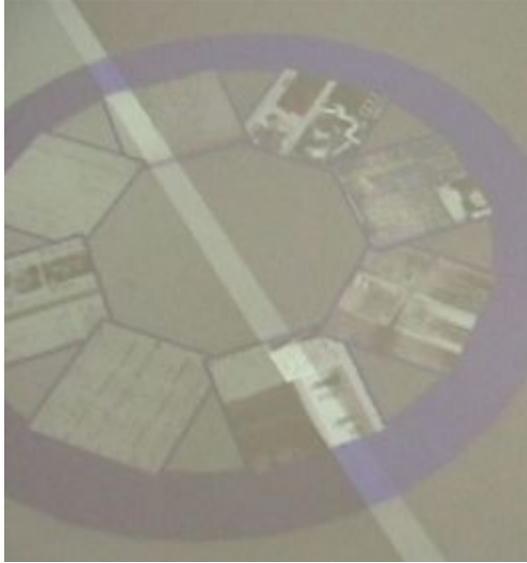


Fig. 6. A shared workspace. Each user can download shared files displayed in the shared workspace by drag-and-dropping these icons onto his/her personal workspace. The user can also upload his/her personal file to this shared workspace by using his/her hand. The user can also rotate this shared workspace by his/her finger

The shared workspace also plays a role as a presentation screen. When a user drags-and-drops any image file on the preview icon, that file is magnified at the shared workspace and all users can see and rotate the image (Fig. 7).

When we use the traditional presentation screen, if one user wants to point to the presentation slide, he/she needs to stand up and go to the presentation screen. On the other hand, the users of our system can point to the slide without leaving their seats. However, all users cannot see the slide from the right angle simultaneously.

3.5 Finger Recognition

To detect hand regions, the system calculates the difference between the initial image and the current image of the tabletop. After the binarization operation, objects which are larger than a threshold are recognized as hand regions. Since the hand regions must cross the edge of the image, the system investigates every pixel of the four edges and determines the direction of finger tips.

After the hand regions are detected, a certain area from the fingertips is recognized as the palm. By repeating the shrink operation to the palm region, the final pixel is determined to be the center of the palm. Then, using template matching with a circle template, the system finds the fingertips.

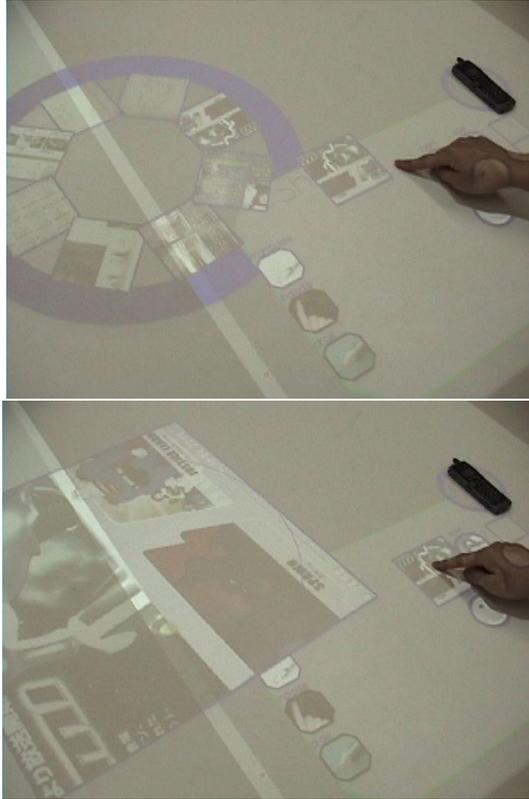


Fig. 7. The user can preview text or image files on the shared workspace by putting the icon on the “Preview” icon

Currently the system can detect and track eight hands in 20 frame/sec. If the CPU becomes much faster, the system can track more hands faster.

After detecting the hands on the table, it is necessary to decide to whom each hand belongs. To do this, the system currently uses the distance between the identified object and the center of gravity of the hand region and the center of the palm.

1. If the center of gravity of hand region is in the personal workspace, the hand is recognized as the owner of the personal workspace.
2. If the center of palm is in the personal workspace, the hand is recognized as the owner of the personal workspace.

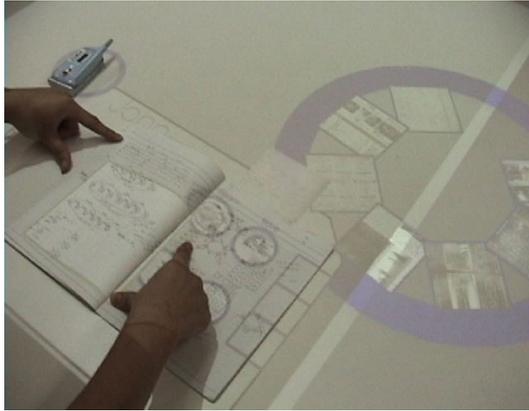


Fig. 8. Interactive image capturing. When the user makes a rectangle gesture with the thumb and index finger of both hands and waits for three seconds, the area is digitally captured and saved

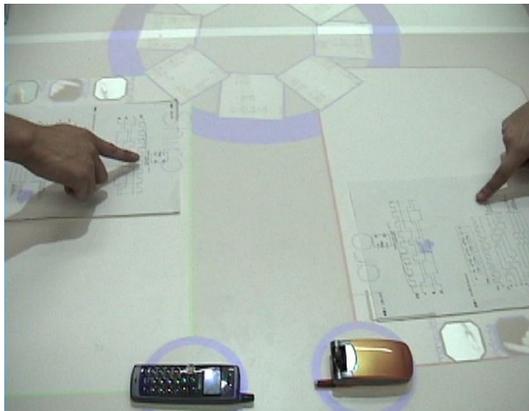


Fig. 9. When the finger sharing mode is enabled, other users' finger positions are projected onto the desktop. If each user aligns the paper document on the grid which is also projected on the personal workspace, the user can see what other users are focusing on

The reason why we need the first rule is that the center of the palm goes outside of the personal workspace when the user moves his/her hand to the shared workspace.

3.6 Interactive Image Capturing

In the current meetings, we sometimes want to digitize a part of or all of the paper document and send it to all participants during the meeting. In order to

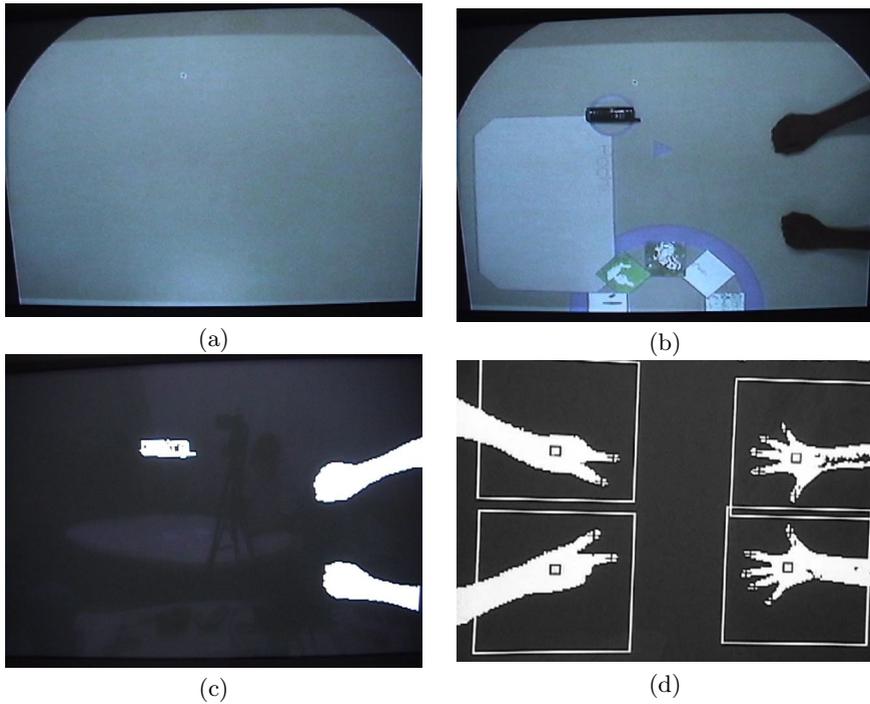


Fig. 10. Hand detection. (a) An initial image of the table. (b) The image when the user's hands and the object are on the table. (c) The difference between (a) and (b). The projected image can be eliminated in this calculation. (d) Four hands are successfully detected and recognized

do this in the current meetings, we need to go to a PC connected to an image scanner, scan the file, and send it as an attachment file. It is a time-consuming task.

On the other hand, our system provides interactive image capturing. When the user makes a rectangle with the thumb and pointing (index) finger of both hands and waits for three seconds, the CCD camera on the ceiling zooms into the rectangle area, and saves its image in JPEG format (Fig. 8). The reason why we used the third CCD camera is to obtain high resolution images when capturing the image.

3.7 Finger Position Sharing

When a user enables the finger position sharing mode, a grid is shown in the personal workspace. If the user aligns a paper document on the grid, other users' finger positions are projected on the document (Fig. 9). Different colors are assigned to each user to identify whose finger position they belong to.

4 Discussion

We have not yet done any formal user studies. It is rather difficult to evaluate the system by comparing it with any existing systems since some elemental technologies, such as user identification, which are essential to our system are not reliable enough. However, through the experimental use in our laboratory and the demonstrations to the visitors, we found that most of the people want to use the system in the real meetings.

The first reason is its natural and intuitive interface for exchanging digital files. In particular, our users preferred to use the virtual Chinese table. However, some people claimed that it is unnatural to use fingers during the meeting because they usually hold a pen. So it might be better to use pen instead of finger to manipulate digital objects, such as seen with tablet PCs. In our framework, it is relatively easy to use a pen instead of a finger to manipulate objects projected on the desk just by exchanging the template image.

The second reason is that people felt much closer to each other during the meeting because there is no physical fence or barrier (i.e., LCD displays of the laptop PCs) between users. In the current meetings, we do not know what other users are doing because of these barriers. Although this is nice to keep our privacy, it sometimes interferes with our mutual understanding.

4.1 User Identification

We used color information of the mobile phones to identify users. This object recognition system gave 92 % correct answers for a hundred different images in our experiment. If the number of users is small and they use phones with different appearances, this identification works. However, if the number of users is larger and they use phones with similar appearance, this identification system would fail.

For the better identification, there are some alternative methods.

1. face recognition
2. RFID tag

The former method is already practically used in some applications. This method works fine when the system can get almost full face of the user in high resolution. However, the image captured by a relatively remote camera would not be good. The latter method is accurate for detecting who is in the environment. However, it is not appropriate for detecting where the user sits. One solution would be to use an object sensor behind the table to detect the precise position of the user.

5 Related Work

There are several works which should be mentioned related to interactions in ubiquitous environment. “Office of the future” [14] proposed an augmented office with some projectors. The users can see information on the desk or on the wall of the office. However, the developers did not mention how to interact in such a

ubiquitous environment. “EasyLiving”[2] proposed an intelligent living room in the future. The environment always tracks users in the room and displays the users’ desktop at appropriate screens. The interaction, however, is done using the traditional mouse and keyboard.

Our work is mostly inspired by Rekimoto’s Augmented Surface[10], which smoothly integrates the desktop of laptop PCs and augmented table surface and wall. Users do not need to know the IP address of computers to exchange digital information with each other. The concept of personal workspace and shared workspace is similar to ours. Augmented Surface is, however, a design prototype based on the current framework. It assumes an environment where people carry their own PCs. On the other hand, our system is a design prototype based on the ubiquitous environment of the near future. In particular, the ubiquitous desktop is a unique concept. Augmented Surface’s vision system, called DeskSat, takes 30 seconds to scan the table surface. On the other hand, our vision system processes eight hands and other object in 20 frame/sec (i.e. 0.05 sec).

Another related work which should be mentioned is the “I/O bulb” concept in [18]. Unlike the normal bulb which has just the ON/OFF state, the “I/O bulb” has the ability of data projection and image capturing. As we described previously, EnhancedTable consists of two unit which includes a LCD, a CCD camera, and two PCs. This unit is a concrete implementation of the I/O bulb concept. We carefully designed our hardware and software so that each unit works by itself and is extensible. As a result, it is easy to add other units to the system in order to make a larger table (or room).

There are many systems that project digital information on the desk. The pioneering work of an augmented desk interface was done in Digital Desk [21]. Digital Desk proposed a basic hardware setup which has been used in later research. Digital Desk also experimented with basic finger recognition. Kruger [6] and MacKay [8] also experimented with augmented desk systems. MetaDesk [4] used real objects (Phicons) to manipulate digital information such as electronic maps. These systems are designed to be used by one user.

PDH[19, 13] is a circular table top display system. It is designed to be used by multiple users from any direction of the circular table. However, it is equipped with a touch panel and cannot detect simultaneous pointing by users. Recently, the touch panels which allow simultaneous, multi-user interaction have been developed (e.g., [3]). They still do not allow users to put some objects. It is also difficult or expensive to make much larger table.

6 Conclusion

This paper described our design and implementation of an augmented table system, called EnhancedTable. The system was designed to be used in the ubiquitous environment in the near future, where people do not need to carry laptop PCs. We proposed and implemented the ubiquitous desktop as personal workspace and the virtual Chinese table as shared workspace. Users can easily exchange their files via the shared workspace in a natural and intuitive man-

ner. We also implemented interactive image capturing and finger point sharing capabilities.

We are currently interested in integrating mobile phones to our system. The mobile phones are used not only for identification but also for browsing files in the personal desktop. We are also interested in developing an intelligent smaller unit which automatically calibrates its camera and its LCD.

References

1. Arai, T., Machii, K, and Kuzunuki, S.: Retrieving electronic documents with real-world objects on InteractiveDesk. In Proceedings of the ACM Symposium on User Interface Software and Technology (UIST'95), ACM (1995) 37–38
2. Brumitt, B., Shafer, S.: Better Living Through Geometry. Personal and Ubiquitous Computing, Vol.5, Issue 1 (2001)
3. Dietz, P.H., Leigh, D.L.: DiamondTouch: A Multi-User Touch Technology, ACM Symposium on User Interface Software and Technology (UIST) , ACM (2001), 219-226
4. Ishii, H., Ullmer, B.: Tangible bits: Towards seamless interface between people, bits and atoms. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'97), ACM (1997) 234–241
5. Koike, H., Sato, Y., Kobayashi, Y.: Interactive Textbook and Interactive Venn Diagram: Natural and Intuitive Interfaces on Augmented Desk System. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2000), ACM (2000) 121-128
6. Kruger, M.: Artificial Reality. 2nd edn. Addison-Wesley, (1991)
7. Lange, B.M., Jones, M.A., Meyers, J.L.: Insight Lab: An Immersive Team Environment Linking Paper, Display, and Data. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'98), ACM (1998) 550-557
8. MacKay, W.: Augmenting reality: Adding computational dimensions to paper. CACM, ACM (1993), Vol.36 No.7, 96–97
9. Nishi, T., Sato, Y., Koike, H: SnapLink: Interactive Object Registration and Recognition for Augmented Desk Interface. In Proceedings on IFIP INTERACT'01, (2001) 240-246
10. Rekimoto, J., Saitoh, M.: Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'99), ACM (1999) 378-385
11. Robinson, P.: Animated paper documents. In Proceedings of HCI'97(21B) (1997) 655–658
12. Sato, Y., Kobayashi, Y., Koike, H: Fast Tracking of Hands and Fingertips in Infrared Images for Augmented Desk Interface. In Proceedings of IEEE Face and Gesture Recognition (FG 2000), IEEE (2000) 462-467
13. Shen, C., Lesh, N., Bardsley, R.S.: Personal Digital Historian: User Interface Design. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2001), ACM (2001) 29-30
14. Raskar, R., Welch, G., Cutts, M, Lake, A, Stesin, L, Fuchs, H.: The office of the future: a unified approach to image-based modeling and spatially immersive displays, In Proceedings of the 25th annual conference on Computer Graphics and interactive techniques, ACM (1998)

15. Streitz, N.A. et al.: i-LAND:An Interactive Landscape for Creativity and Innovation. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'99), ACM (1999) 120-127
16. Streitz, N.A., et al.: Roomware - The Second Generation. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2002), ACM (2002) 506-507
17. Turk, M.: Moving from GUIs to PUIs. In Proceedings of Fourth Symposium on Intelligent Information Media, (1998)
18. Underkoffler, J., Ishii, H.: Illuminating light: An optical design tool with a luminous-tangible interface. In Proceedings of the ACM Conference on Human Factors in Computing System (CHI'98), ACM (1998) 542-549
19. Vernier, F., Lesh, N. Shen, C.: Visualization techniques for circular tabletop interfaces. In Proceedings of Advanced Visual Interfaces (AVI 2002), (2002) 257-265
20. Weiser, M.: The Computer for the 21st Century. Scientific American, Sep. 1991.
21. Wellner, P.: Interacting with Paper on the DigitalDesk. CACM, ACM (1993) 86-96
22. Yamashita, J. Kuzuoka, H., et al.: Agora:Supporting Multi-participant Telecollaboration. In Proceedings of HCI 1999, (1999) 543-547