

An Automated Meeting Assistant: A Tangible Mixed Reality Interface for the AMIDA Automatic Content Linking Device

Jochen Ehnes

Centre for Speech Technology Research, University of Edinburgh
10 Crichton Street, Edinburgh, U.K.
jehnes@inf.ed.ac.uk

Abstract. We describe our approach to support ongoing meetings with an automated meeting assistant. The system based on the AMIDA Content Linking Device aims at providing relevant documents used in previous meetings for the ongoing meeting based on automatic speech recognition. Once the content linking device finds documents linked to a discussion about a similar subject in a previous meeting, it assumes they may be relevant for the current discussion as well. We believe that the way these documents are offered to the meeting participants is equally important as the way they are found. We developed a mixed reality, projection based user interface that lets the documents appear on the table tops in front of the meeting participants. They can hand them over to others or bring them onto the shared projection screen easily if they consider them relevant. Yet, irrelevant documents don't draw too much attention from the discussion. In this paper we describe the concept and implementation of this user interface and provide some preliminary results.

Keywords: Meeting assistants, Meeting processing, Mixed reality(MR), Projected user interface, Tangible user interface (TUI).

1 Introduction

While the main purpose of meetings is to facilitate direct communication between participants, documents play an important role in meetings as well. Documents often contain facts that are currently discussed, but they are not necessarily at hand. If these documents were available in a document management system, participants could search for them. However, participants of a meeting usually do not have the time to perform such queries often during a meeting. Therefore a system that could provide relevant documents for an ongoing discussion would be very helpful. A critical part of such a system would be the user interface. It should stay in the background as much as possible in order to not disturb the ongoing discussion by drawing too much attention to it. Yet it should be able to deliver the relevant documents to the participants as directly as possible, so they can incorporate these documents into the discussion directly with minimal effort. In this paper we describe a tangible mixed reality system as an interface for the AMIDA Content Linking Device [7], a system that suggests documents which may be of interest for an ongoing discussion. The documents suggested by this content linking

device were displayed on a laptop screen. Consequently a meeting participant, usually the discussion leader, has to monitor what is going on on the laptop's screen, which certainly distracts him or her from the meeting. Furthermore the laptop's display has the character of a private display. Other participants are not able to see the documents although they may be more important to them.

Of course this could be fixed easily by providing every participant with a laptop showing all the proposed documents. However, then everybody would have to check every document the system suggests. Furthermore, if a participant thinks a document is important, it still would not be straight forward to share it with the other participants. The user would have to describe the document first, so that the others can identify it among all the documents the system suggested so far. All this would lead the attention too much onto the laptop in front of the participants and away from the group. To overcome these challenges we propose to use a user interface projected on the tabletop in front of the participants. By using this less private form of display the documents suggested by the content linking device are visible to other participants as well. By furthermore providing an easy way to grab these documents and move them to other participants' places quickly, they can be moved to the participant they are most valuable to. Documents that are of interest to several participants or that are the subject of the discussion can be moved to a shared space, where they can be looked at by everybody at the same time.

2 Previous and Related Work

Since January 2004, the European AMI (Augmented Multi-party Interaction) integrated project has been building systems to enhance the way meetings are run and documented. AMI research revolves around instrumented meeting rooms which enable the collection, annotation, structuring, and browsing of multimodal meeting recordings. AMIs JFerret browser [4] allows its users to go through previous meetings to get themselves up to date if they were not able to attend these meetings. The browser can display video and audio recordings of all participants as well as the transcript of what was said by whom. Searching for keywords makes it easier to find parts of particular interest. While the possibility to look through recordings of previous meetings and being able to search for important sections by keywords is a very helpful tool, it requires direct action by the user. Furthermore, as the user has to interact with the system on a personal computer, which draws the user's attention to it and distracts from the conversation, the browser is more useful in preparation for a meeting than during the meeting itself.

An important goal of AMI however is to support meetings while they take place. An automated meeting assistant shall find relevant information based on the current discussion and offer it to meeting participants without requiring too much attention from them. The AMIDA Content Linking Device [7] is the first demonstrator of this system. The system consists of a Document Bank Creator (DBC) that gathers documents that are of potential interest for an upcoming meeting, a Document Indexer (DI) that creates an index over the document bank created by the DBC, a Query Aggregator and a User Interface, all connected via a central Hub. During a meeting, the Query Aggregator performs document searches at regular time intervals using words and terms from the

automatic speech recognition system. It produces a list of document names, ordered by relevance, based on the search results, as well as on a persistence model, assuming that documents that come up during several searches are likely to be more relevant than others that do not. The User interface finally displays the results to the user(s). As this work is about an alternative User Interface, we can't go into more details of the content linking device and refer to [7] for that.

The idea to use the table top as an interface to computers is not new. The first system of that kind known to the authors was DigitalDesk [12,13]. Its main intention was to bring together electronic- and paper-documents.

In [10,11] a similar setup consisting of video projector and camera, I/O-Bulb as the authors call it, mounted above the table was used to create applications that are manipulated using physical objects. Applications include the simulation of holography setups using physical placeholders for optical elements such as lasers, mirrors etc. or the simulation of fluids flowing around objects standing on the tabletop. An obvious advantage of this kind of user interface is their collaborative nature as several users can manipulate different physical objects on the tabletop at the same time, instead of being restricted to a single mouse in a conventional computer setup.

While being able to see what everybody else sees is a very important factor for collaboration, it sometimes is necessary to be able to sketch something down to clear one's thoughts before presenting it to the whole group. In [3] the authors presented a system that supports the discussion of virtual prototypes by a group of designers/engineers sitting around a projection table. The crucial difference here to other 3D viewers or the applications running on the I/O-Bulbs was that the content does not occupy the whole screen space. Instead the virtual prototype would be visible on a virtual piece of paper. Apparently conventional plots of CAD drawings were still used frequently during design review meetings, as it was so easy to sketch on them with a pencil to point out flaws or suggest improvements. Furthermore, one could just work out an idea on a piece of paper before presenting it to other meeting participants. The computer systems at the time were just considered a hinderance during these meetings. In order to make them more usable for such applications, the above mentioned prototype was developed. As the 3D models were displayed on the virtual pieces of paper, they were visible to everyone. Furthermore, the paper could be moved around using a tracked puck, so that they could be brought closer and rotated to a single person to allow for a more personal use. By grabbing two points, one with the puck and one with the pen, the virtual paper could be scaled, similarly to the two finger pinch scaling gesture known from the iPhone. Using tracked pens, participants could draw lines on the objects to annotate them. Furthermore, the system allowed to connect each piece of paper to one of several tracked pairs of shutter glasses to get a full three dimensional impression of the object. But as the stereo view certainly hindered others looking at the object, it could easily be switched off again by putting the glasses down for discussions. While we do not display 3D objects in our content linking system, we use the concept of having virtual pieces of paper that can be moved around using a physical device such as the puck. The Shared Design Space [5], a system consisting of four projectors for an interactive surface on a tabletop and a projector to create an interactive wall, is of interest as it not only use video cameras to track objects for interaction. Anoto Pens, digital pens that can track a

pattern of tiny dots on the surface they are writing on, are used to control the system as well as to draw onto the virtual documents (images). As we aimed for a simple interface to view existing documents, we don't provide such a feature at the moment.

3 Setup

In order to present documents found by the content linking device to meeting participants, we planned to build a system that can project these documents on the table tops in front of the participants during meetings.

As one of our goals for the implementation of this system was to provide the additional functionality without requiring big changes to the existing meeting environment, we decided to go for projection and video tracking from above the table. While a back-projection/tracking system from below a semi transparent tabletop would have allowed to detect when objects or fingers touch the surface, having to buy new tables with transparent tops and fitting back projection systems beneath them would have been a too big a change to the existing room as it was already fitted with a lot of recording equipment used by the AMI project. New furniture might have made other changes necessary, which we wanted to avoid to keep recordings comparable. Furthermore, the space required for a back projection/tracking system below the table possibly would become a disadvantage as meeting participants would not have been able to sit at the table as comfortably as usual and projected documents may come and go unnoticed if papers or other objects are put on the desk top on top of them.

In order to provide enough space for several participants, we planned to use multiple projection/tracking systems. To start we designed our system to support two users and we used one computer (Mac mini), projector (Optoma EP709) and camera (Imaging-Source DFK 21BF04) for each of them. Figure 1(a) shows our setup mounted around the projector for the presentation screen. We also included the presentation screen as a shared space into our system, by feeding this projector with a laptop computer (MacBook Pro).

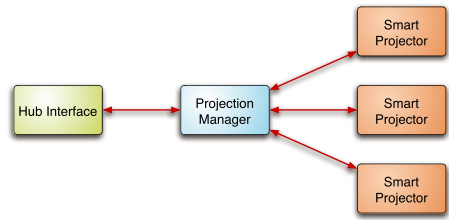
However, the software is designed to be scalable, so that we would be able to change the number of projection systems easily. While multiple projectors are often used to create a larger, tiled display [8,6], that approach did not appear to be so suitable for our application. Desks in meeting rooms often are not arranged to create one large surface, but in different shapes, such as a "U" shape, to allow everybody to see everyone else as well as a projection screen. Some tables may not even be connected to others at all. In consequence it is not important that the projection systems form a consistent display area, as long as the user interface is consistent across the systems and documents can be moved between them in a way that is consistent to the way they are moved around on one system. Furthermore, this approach gave us the flexibility to provide individual display modes for each user, such as the shared display layer described in section 5.3.

4 System Architecture

In order to keep the number of projection systems scalable, we divided our projection system into two parts, Smart Projector an application running on all projection units and



(a) The two projection systems Mounted around the existing projector.



(b) System architecture.

Fig. 1. Prototype setup and System Architecture

a central Projection Manager. This display system is connected to the Content Linking Device via a third application, the Hub Manager (figure 1(b)).

4.1 Smart Projector

Smart Projector is the application running on every projection unit that creates the actual user interface for meeting participants. While It has a simple user interface to connect to / disconnect from the Projection Manager and to configure and activate the capturing of life video from the camera, it is switched to Fullscreen mode during normal operation. If a video input stream is available, the application searches for AR-ToolKit+ markers in it and sends the information to the Projection Manager and the applications associated with the tracked object. If no video stream is available, it displays shared documents (section 5.3) only. Additionally it also captures key board events and forwards them accordingly (section 5.2).

4.2 Projection Manager

At the center of this projection system is Projection Manager, a server that manages all important parameters of the projection units and coordinates their actions. As the Mac minis only have one display connector feeding the projector, most parameters of Smart Projector (background color, projection parameters, calibration between projection and camera coordinates, ...) are adjusted in the Projection Manager. This also makes it easy to adjust parameters for several projection units at once.

Besides managing the projection units, Projection Manager is also used to define and print the interaction devices carrying AR-ToolKit+ markers (see figure 2(b) for example) and to manage the display applications running on the system.

4.3 Display Applications

In order to keep the system easily extensible to new types of documents, we developed an API based on two base classes that can be extended to create different display

applications. A peculiarity of that API is that it consists of two base classes, one for application objects and one for display objects. The display objects are basically stateless objects that render the content they are sent by the application object on table top in the form of a sheet of paper. They are also responsible to forward certain events to the application object. The application objects on the other hand are responsible to maintain the state and change it according to user input. Whenever relevant parameters change, the application objects have to send updates to their display objects. This separation of state and display allows for an easy duplication of the display. When a display application is started on one system and is not running anywhere else, a display object as well as an application object is created locally on the projection unit. On the other hand, if the application is already running somewhere else, for example if it has been moved to the shared space (section 5.3) and consequently it has to be started on other machines, only a display object is created. This display object is then connected to the application object on the projection unit where the application was started first. After that the application object receives a call to update its display objects, so that the newly created display object displays the correct data.

4.4 Hub Interface

To connect the display system to the AMI Content Linking Device the Application Hub-Interface was developed. On one side it connects to the Hub as a consumer via the Java Native Interface, and on the other side it connects to the Projection Manager. Once the query aggregator stores new related documents in the Hub (please refer to [7] for a detailed description of the components of the content linking device), the application receives a message from the hub containing the documents name.

Upon receiving this message, the Hub Interface introduces these documents, which were all converted to PDF, into the projection system. It does so by adding a new PDF-Reader application to the list of applications and setting the Document URL as well as the Application ID to the URL of the PDF file.

Then the application is started on a projection unit specified by a popup menu. In a future version the content linking system should also provide a person or role for whom the document is most relevant. Then the documents can be sent to the best fitting person automatically. If the document is already being displayed, it is not introduced a second time. Instead the user is notified that the document could be relevant for the current discussion by bringing it to the front and letting it vibrate a little bit to create a visual ping (section 5.4).

5 User Interface

In order to make the interaction with the system as direct as possible, we aimed to make the projected objects graspable. We decided against hand tracking as it is difficult to distinguish between gestures meant to manipulate documents and gesturing during discussions. This is especially true as the current setup does not allow to detect if the user's hands touch the desk top. Instead we track physical objects that serve as interaction devices using the AR-Toolkit+ tracking library.

5.1 Document Handling

In order to move projected documents around, a physical object (paper grabber) is associated with them. As long as this connection exists, the virtual document follows the grabber.

The grabber objects consist of a piece of cardboard containing three markers, one of them elevated on a box (see figures 2(b) and 2(a)).

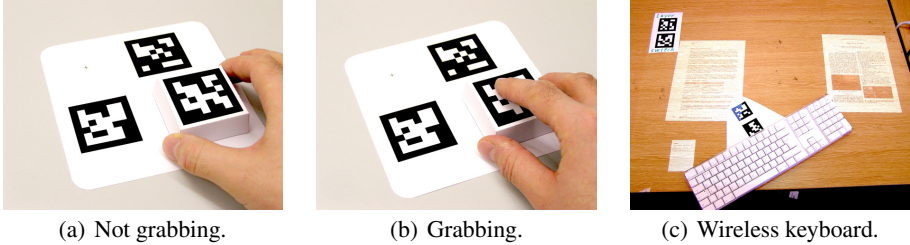


Fig. 2. Input devices

The elevated marker has the functionality of a switch. By blocking the visibility of this marker using a finger for example, one tells the system to grab the document below the device. If grabbed only on the sides, so that the marker on top is fully visible, the marker is disconnected and can be moved freely. Once the marker is placed on a virtual piece of paper, users can grab the document by holding the box like a mouse and thereby covering the top marker. Because the switch marker is on the box that user's grab, they do not have to think consciously about covering the marker or not. They just have to remember the two ways of holding the grabber device; On the sides to lift the grabber from the paper or with the hand on top of it, pressing it onto the paper they want to move. Once grabbed, the document stays connected to the grabber until it is released again, i.e. the top marker gets recognized again. This may be on the same or another user's projection system.

5.2 Keyboard Forwarding

Instead of providing virtual, projected keyboards as it is usually done with touch screen interfaces, we chose to use standard wireless keyboards. In order to allow keyboard based input, a Keyboard identified by the two markers attached to it (see figure 2(c)), can be placed on a displayed document. This allows to route keyboard events to the display applications (section 4.3) that create the graphical representation of the documents. It replaces the physical connection (which keyboard is connected to which projection unit) with a virtual connection between keyboards and documents.

5.3 Sharing

In addition to augmenting the table, we wanted the system to incorporate the whiteboard as well. This way, participants are able to interact with content on the whiteboard directly from their place and move content between their space and the shared whiteboard space

easily. While Hyperdragging as described in [9] would allow participants to do that in principle, it relies on a laptop with a conventional interaction device such as a touchpad. Using hyperdragging therefore would work against our goal to let the computer disappear.

We believe it is better to 'bring the shared screen to the participant' on the press of a button, or in our case when a marker is covered by the user. We therefore implemented a shared semitransparent layer (see figure 3(a)) on top of the normal projection area which can be activated and deactivated by covering a marker placed on the projection area for that purpose. The presentation screen is the only exception here, as it does not have a private layer. It always displays the shared layer. Documents can be moved between the private and shared layers by grabbing them on one layer before switching to the other one. Once on the shared layer, all state changes such as position, orientation or which page of a multi page document is shown are forwarded immediately to other systems displaying the shared layer.

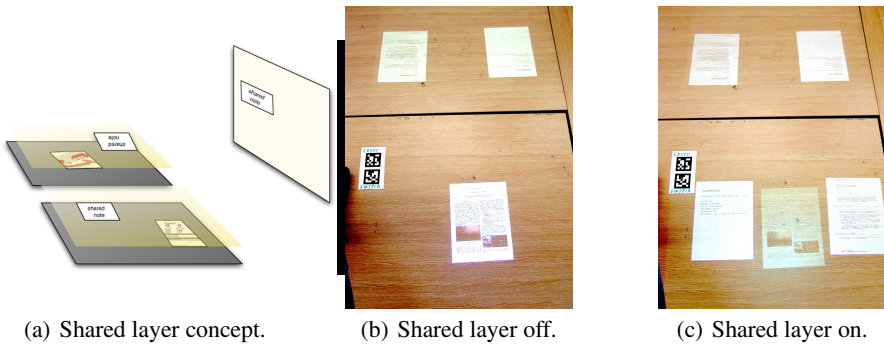


Fig. 3. Shared space

5.4 Auto-arrangement and Auto-iconizing

The Content Linking Device brings up new documents in regular intervals. In fact it often finds several documents to be displayed at the same time. In this situation it obviously is not enough to make the documents appear at a fixed location (eg. the center of the tabletop). Of course the space on the table is not unlimited, so a method had to be developed to prioritize documents and remove less relevant documents gradually. We implemented a system to arrange and iconize documents automatically. It behaves as follows: When a display application is started to present a document, it is appended to the array of automatically arranged applications. If the number of elements in this array is growing above a given limit (two applications in our case), the first element is removed and appended to the array of iconized applications. Additionally, a timer is set for each application added to the array of automatically arranged applications. Once the timer fires, the document gets iconized as well. This way documents that do not appear to be relevant to users are removed as well. If the number of elements in the array of iconized applications grows above its limit (ten applications in our case), the first application is removed and terminated. Whenever applications are added to or removed from these arrays, the applications are sent a new goal position and scale factor according to

the array and their position within that array. The first auto arranged application is displayed on the left side. The second (and latest) one is positioned next to it in the center of the projection, leaving the right side for documents the user places there to read.

Their scale factor is 1.0 so they are displayed in full size. The automatically iconized applications on the other hand are scaled down to 0.3 and arranged along the front edge of the table with the oldest one being displayed on the left and new ones being added to the right. When applications are sent to new positions or receive new scale factors, they don't change to these values immediately. Instead they animate towards these values over a given duration (1.5 seconds seemed best). This way it becomes obvious when the layout changes and it is easy to follow what is going on. This is very useful when an application that is already open is deemed to be relevant by the query aggregator again, as one can see the document move from its previous position to the position of the newest document (center). If the user places a paper grabber or keyboard connector on top of a virtual paper, it prevents the paper from being affected by the auto arrangement/iconizing system. If placed on an iconized paper, the paper is also scaled up to full size again. Now the user may move the document to where it can be read conveniently without interference of the auto arrangement system. Once the user removes the paper handling device from the projected document and no keyboard is connected with it, the system will take responsibility for it again and iconize it after 30 seconds to clean up the tabletop. For the case that the document the query aggregator determined as relevant is already displayed as the latest document or controlled by the user, a visual ping has been implemented. If pinged, a document visually vibrates for a short period of time. It is animated by scaling it slightly up and down from its original size using a sine function. The amplitude of this vibration is scaled down to zero within 1.5 seconds to fade out the effect smoothly.

6 Results

We developed a scalable projection system to be used in meeting environments. The way it is set up allows for easy installation in existing environments. After all, the camera, projector and computer can be integrated into a single unit mounted above the tables. We implemented software components that allow for easy management and coordination of the projection units as well as a user interface based on tracked interaction devices. We demonstrated that it is easily possible to move documents around on one, as well as between different projection units or between private and shared spaces. Furthermore, the system is able to connect to the central Hub of the AMI project. This way it can be used to display documents the content linking device deems relevant for the ongoing discussion. Additional functionality to manage the displayed content automatically was implemented to cope with the stream of new documents being introduced by the content linking device.

As it turned out, the first version of the content linking device we used in our prototype had not been tuned well enough yet. During the trial meeting¹ it repeatedly brought up two agendas but rarely anything specific.

¹ Meeting ES2008d of the AMI Meeting Corpus [2]. For testing purposes meeting ES2008d is played back as input to the Content Linking Device, which searches for relevant documents from the meetings ES2008a/b/c. For more details please refer to [7].

Another weakness of the current system is the resolution of its projectors (1024x768). While this is ok for documents containing little text in a large font, such a meeting agenda or Power Point slides, it is not sufficient to display regular text documents as a virtual sheet of A4 paper.

Mark Ashdown [1] proposed to use two projectors, one to cover the whole table top with a relatively low resolution and one to cover a small area with a high resolution. However, we believe that this would effectively limit the usable display area to the small area of high resolution projection and as such would have a negative effect on our system. Participants would have to move all documents onto this foveal display area to be able to read them, which would introduce additional load to the user. It would no longer be possible to just throw a glimpse at a new document. Furthermore, it would also make it impossible to place a virtual paper between two participants to look at it together, as that would be in the low resolution area.

7 Future Work

In order to address the readability problem we plan to use a projector capable of projecting full HD video (1920x1080). We furthermore plan to use it in portrait mode, effectively augmenting only one half of the users' table space. This increases the resolution of the display area further and as documents are usually printed in portrait format, it should enable us to make better use of the projected pixels.

We may also look into different approaches, such as the Microsoft Surface Computers. On the one hand they still have the disadvantages of back projection systems as described in section 3. For example they cannot be integrated into an existing setup as easily as an on-projection system and maybe more importantly, they do not allow their users to put their legs beneath the table to sit comfortably for extended periods of time. On the other hand they come as a complete unit which is easy to set up and they appear promising, as they go beyond the usual touch screens and allow to recognize objects laid onto them. This feature is necessary to attach virtual documents to physical (grabber) objects, which can easily be handed around between different meeting participants, in order to move documents between different units. However, as the resolution of the current surface computers is 1024x768, we expect them to have the same limitations as our current projection system when it comes to displaying text documents with standard sized print.

Once the display quality is sufficient for text documents and the content linking device gets better tuned, we plan to include the system in the scenarios for future AMI meeting recordings. This should give us the possibility to evaluate the system in a formal way.

Acknowledgements. This work is supported by the European IST Programme Project FP6-0033812 (AMIDA) as well as the Marie Curie Intra European Fellowship (IEF) FP7-221125 (NIPUI). This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

References

1. Ashdown, M., Robinson, P.: A personal projected display. In: MULTIMEDIA 2004: Proceedings of the 12th annual ACM international conference on Multimedia, pp. 932–933. ACM Press, New York (2004)
2. Carletta, J., Ashby, S., Bourban, S., Flynn, M., Guillemot, M., Hain, T., Kadlec, J., Karaiskos, V., Kraaij, W., Kronenthal, M., Lathoud, G., Lincoln, M., Lisowska, A., McCowan, I., Post, W., Reidsma, D., Wellner, P.: The ami meeting corpus: A pre-announcement. In: Renals, S., Bengio, S. (eds.) MLMI 2005. LNCS, vol. 3869, pp. 28–39. Springer, Heidelberg (2006)
3. Ehnes, J., Knöpfle, C., Unbescheiden, M.: The pen and paper paradigm supporting multiple users on the virtual table. In: Proceedings of the Virtual Reality 2001 Conference (VR 2001), p. 157. IEEE Computer Society Press, Los Alamitos (2001)
4. Fapso, M., Schwarz, P., Szöke, I., Smrz, P., Schwarz, M., Cernocký, J., Karafiát, M., Burget, L.: Search engine for information retrieval from speech records. In: Proceedings of the Third International Seminar on Computer Treatment of Slavic and East European Languages, pp. 100–101 (2006)
5. Haller, M., Brandl, P., Leithinger, D., Leitner, J., Seifried, T., Billinghamurst, M.: Shared design space: Sketching ideas using digital pens and a large augmented tabletop setup. In: Pan, Z., Cheok, D.A.D., Haller, M., Lau, R., Saito, H., Liang, R. (eds.) ICAT 2006. LNCS, vol. 4282, pp. 185–196. Springer, Heidelberg (2006)
6. heyewall. Heyewall, <http://www.heyewall.de/>
7. Popescu-Belis, A., Boertjes, E., Kilgour, J., Poller, P., Castronovo, S., Wilson, T., Jaimes, A., Carletta, J.: The amida automatic content linking device: Just-in-time document retrieval in meetings. In: Popescu-Belis, A., Stiefelhagen, R. (eds.) MLMI 2008. LNCS, vol. 5237, pp. 272–283. Springer, Heidelberg (2008)
8. Raskar, R., Welch, G., Fuchs, H.: Seamless projection overlaps using image warping and intensity blending. In: Proceedings of the Fourth International Conference on Virtual Systems and Multimedia, Gifu, Japan (November 1998)
9. Rekimoto, J., Saitoh, M.: Augmented surfaces: A spatially continuous workspace for hybrid computing environments. In: Proceedings of CHI 1999 (1999)
10. Underkoffler, J., Ishii, H.: Illuminating light: An optical design tool with a luminous-tangible interface. In: CHI, pp. 542–549 (1998)
11. Underkoffler, J., Ullmer, B., Ishii, H.: Emancipated pixels: real-world graphics in the luminous room. In: Rockwood, A. (ed.) Proceedings of the 26th annual conference on Computer graphics and interactive techniques, pp. 385–392. ACM Press/Addison-Wesley Publishing Co. (1999)
12. Wellner, P.: The digitaldesk calculator: Tangible manipulation on a desk top display. In: Proc. ACM SIGGRAPH Symposium on User Interface Software and Technology, pp. 107–115 (1991)
13. Wellner, P.: Interacting with paper on the DigitalDesk. Communications of the ACM 36(7), 86–97 (1993)