

Diploma Thesis: "Evaluating Collaborative Work Enhanced By Different Augmented Reality Avatar Representations"

Supervising Professors:

- Prof. Grudrun Klinker (Technical University Munich)
- Mark Billinghurst (Hitlab New Zealand)

Supervisor:

- Martin Wagner (Technical University Munich)

Introduction

Desktop video and audio conferences are naturally used today. But there are still some problems left that make it still necessary to travel to clients or colleagues. Some of these are low resolution and bad frame rate of the video stream that reduces the convenience and usability and therefore the acceptance, because important non-verbal information like gaze gets lost. Also the feeling or awareness what the other users are doing or where the user is actually looking is missing. That's because you only have a small picture of the partner and you are not aware of his/her working environment.

The aim of this thesis is to investigate how you can increase this feeling or awareness. During a user study I want to analyze if a live video stream is really necessary or if an avatar representation or a simple picture of the participants could be enough.

Used techniques and theoretical background

Augmented Reality

Augmented Reality is the technique you use to enrich your environment with special information. A common application is to display virtual objects in a live camera picture (e.g. see the Magicbook [6]).

To retrieve more information about the working environment and to visualise this information again as intuitive as possible I want to use the ARToolkit developed by Mark Billinghurst and Hirokazu Kato [5]. The toolkit calculates the position and orientation of special markers in a video stream. This information can be used for tracking of objects or for projecting virtual object in the scenery captured with the camera.

Some CSCW aspects

You cannot imagine today's business life without team or collaborative work. There is a separate branch of research that examines work life and tries to develop supporting tools: CSCW (*Computer Supported Cooperative/Collaborative Work*). The focus is to support team work with the most appropriate techniques. One also often talks of the seamless integration of the computer in

workaday tasks. E.g. that you use your handheld the same way now as you were using your pocket diary.

Baecker [2] introduced a taxonomy to distinguish the different ways of support. He separates in the two dimensions *time* and *space*:

	One meeting site (same places)	Multiple meeting sites (different places)
Synchronous communication (same time)	Face to Face Interactions <ul style="list-style-type: none"> • Public computer displays • Electronic meeting rooms • Group decision support systems 	Remote Interactions <ul style="list-style-type: none"> • Shared view desktop conferencing systems • Desktop conferencing with collaborative editors • Video conferencing • Media spaces
Asynchronous communication (different time)	Ongoing Tasks <ul style="list-style-type: none"> • Team rooms • Group displays • Shift work groupware • Project management 	Communication and Coordination <ul style="list-style-type: none"> • Vanilla email • Asynchronous conferencing bulletin boards • Structured messaging systems • Workflow management • Version control • Meeting schedulers • Cooperative hypertext & organizational memory

We assume here that we only speak of synchronous work. The space dimension is just open. But the main focus will be in distributed collaboration but perhaps some investigation will reveal usability also on a shared space.

DWARF

Because it will be about a distributed application that should provide the maximal flexibility and scalability, I want to build the application on top of DWARF

[4].

DWARF (Distributed Wearable Augmented Reality Framework) is a framework developed at the department of Software Engineering at the Technische Universität München. It is based on CORBA and enables you to easily connect different processes over the network. The main principle behind DWARF are the so called *services*. Every service is a own process and has different *needs* and *abilities*. The services can communicate with others via these needs and abilities that have to match. Between these needs and abilities different communication protocols can be established:

Event based services can subscribe to a channel and receive all events that are send from another service

Object references one service exports an interface that other services can import and call the emthods on the remote object

shared memory two services can easily exchange data between eachother via a shared part of the memory (this feature is currently only available under Linux)

The Setup

The setup should represent a meeting of multiple persons that are sitting around a table and work together on a shared object. That could be a model of a car, a contract or a presentation. We want now to distribute this meeting so that every user can stay at his desk within his working environment but the feeling should be that they are working very close together.

As output device a hand held display or a HMD (head mounted display) is used that is combined with a camera. So the user has the impression that he directly sees through the glasses in the real world that is augmented with the virtual objects.

We assume in the following that there is a conference with three participants (A , B and C).

Every participant has three markers. Two are used for the representation of other participants ($M_{A,B}$, $M_{A,C}$, $M_{B,A}$, $M_{B,C}$, $M_{C,A}$, $M_{C,B}$) and one for the shared workspace ($M_{A,W}$, $M_{B,W}$ and $M_{C,W}$). So if user A is looking through the HMD at Marker $M_{A,B}$ he can see for example the avatar representation of user B .

Every participant can place these marker freely on his/her desktop.

Assume now that person C looks at person B (actually $M_{C,B}$). At person A the virtual representation of person C (on marker $M_{A,C}$) is rotated in direction of person B ($M_{A,B}$). So you can get a rough sense where the other users are looking.

If a user is looking at the workspace marker $M_{x,W}$ as a special feature it should be visualized at the other users where the user is looking on the marker and the virtual object on the marker respectively. This could be done by small arrows that are pointing at the object from the direction of the users marker.

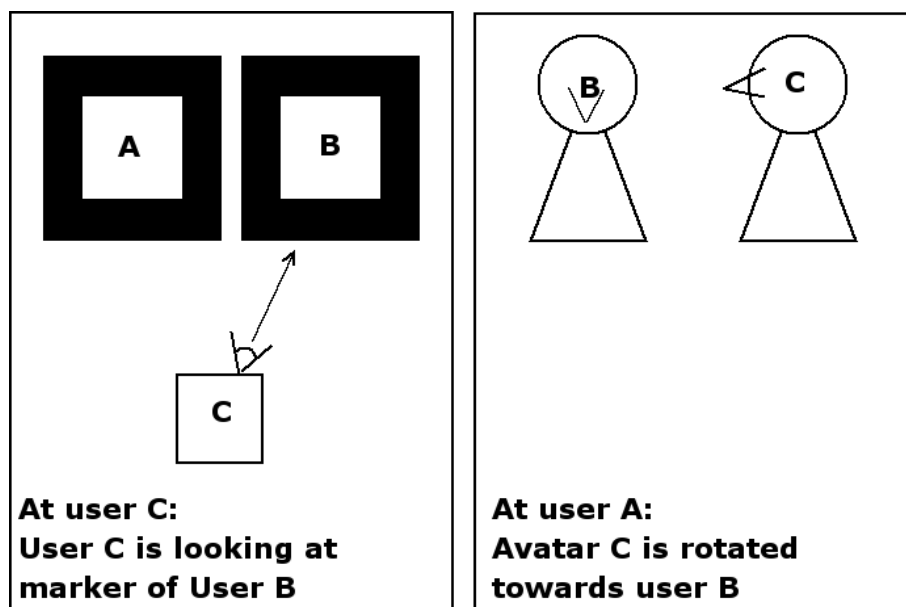


Figure 1:

Different Evolutionary Steps

To fast achieve first results I want to implement a simple avatar representation of the user. In the second step a photo of the user should act as representation and if possible at last step a live video stream shall be displayed instead of the avatar.

This bottom-up procedure will also allow me to investigate if user really have the need for a live video picture or if a simple avatar representation could be enough.

Research has shown that it is not always necessary that there is a video connection [8].

Evaluation

In a final user study I want to test the different representations. The users will get a shared challenge. There are two possible tasks.

1. One would be the Map Task [1] where two user each get a map. In one map there is a route marked. Both maps are similar but have different features. Now the 'information giver' has to explain the route to the 'information follower'. One could modify this task that there are three user and they have to find a route from A to B where in every map different features are missing or there are some white spots.

The interesting thing could be especially that you can test the feature that you get the awareness where the single participants are looking at on the shared marker.

2. Another idea would be that every user has four bricks. And two of these four bricks are equal to all users. The users then have to find out which bricks are the same at every user.

Current research and own focus

There has been different research been done and still going on that hit some aspects of this work:

In this paper [7] an application is described where the video stream is projected on markers that are free to position. But there is no real network connection implemented.

At the Studierstube they investigated how you can augment a video stream of a videoconference. They found that you can not use the video stream for marker detection since the compression destroys the videoframes too much. So they developed a method to send the marker data synchronous to the video data stream [3].

At in the cAR/PE! project at DaimlerChrysler a research group is developing techniques to interact with a shared object during a meeting [9].

In contrast I want to focus to enrich the awareness of the group members. Thus I want to track the users view and develop methods to virtualize these collected informations.

References

- [1] Anderson AH, Garrod SC, Clark A, Boyle E, and Mullin J. The human communication research centre dialogue database. *J Child Lang*, 1992.
- [2] Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg. *Readings in human-computer interaction. Toward the Year 2000*. Morgan Kaufmann Publishers, 2nd edition, 1995.
- [3] Istvan Barakonyi, W. Frieb, and Dieter Schmalstieg. Augmented reality videoconferencing for collaborative work. In *Proc. of the 2nd Hungarian Conference on Computer Graphics and Geometry*, 2003.
- [4] Martin Bauer, Bernd Bruegge, Gudrun Klinker, Asa MacWilliams, Thomas Reicher, Stefan Riss, Christian Sandor, and Martin Wagner. Design of a component-based augmented reality framework. In *Proceedings of the International Symposium on Augmented Reality (ISAR)*, October 2001.
- [5] Mark Billinghurst and Hirokazu Kato. Artoolkit. <http://www.washington.edu/artoolkit/>, 2002.
- [6] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. Magicbook: transitioning between reality and virtuality. In *CHI '01 extended abstracts on Human factors in computing systems*, pages 25–26. ACM Press, 2001.

- [7] H. Kato, M. Billinghamurst, K. Morinaga, and K. Tachibana. The effect of spatial cues in augmented reality video conferencing. In *9th International Conference on Human-Computer Interaction*, pages 478–481, 2001.
- [8] Claire O'Malley, Steve Langton, Anne Andreson, Gwyneth Dotherty-Sneddon, and Vicki Bruce. Comparison of face-to-face and video-mediated interaction. *Interacting with Computers*, 8(2):177–192, 1996.
- [9] Holger T. Regenbrecht and Michael T. Wagner. Interaction in a collaborative augmented reality environment. In *CHI '02 extended abstracts on Human factors in computing systems*, pages 504–505. ACM Press, 2002.