AUGMENTED MULTI-USER COOPERATIVE ENVIRONMENT

Ing. Vítězslav BERAN, Doctoral Degree Programme (2)
Dept. of Computer Graphics, FIT, BUT
E-mail: beranv@fit.vutbr.cz

Ing. Jan PEČIVA, Doctoral Degree Programme (3)
Dept. of Computer Graphics, FIT, BUT
E-mail: peciva@fit.vutbr.cz

Supervised by: Dr. Pavel Zemčík

ABSTRACT

This paper presents a concept of augmented environment system for multi-user cooperation. The system is composed from two main parts: scene recognition and augmentation part and consistent data sharing engine. Scene recognition process is based on common computer vision techniques. Detected objects are shared between all system users keeping the object consistency. Virtual object are than used for environment augmentation.

1 INTRODUCTION

In the age of global cooperation between several distant research groups, such topic becomes more and more important. We propose a concept of sharing real and virtual information using augmented environment. The focus of our work is to recognize the workspace and share it between two remote users preserving the environment consistency. The partial scene recognition is based on common computer vision techniques extended by methods of the augmented reality. Augmented reality (AR) is a technology that attempts to restrain drawbacks of the virtual reality. Instead of the virtual reality techniques, AR does not create any artificial environment around the user but only the user's augmented view. The computer-created virtual objects are inserted into the image of the real world so the authenticity and naturalness are maintained 0.

Collaboration between users requires sharing of the virtual scene objects and data. Virtual environment with shared objects is often called Collaborative Virtual Environment (CVE). CVE methods for sharing objects in this application will be presented in the paper.

2 SYSTEM OVERVIEW

The system structure is motivated by our previous work presented in [2] and it is composed from two separate workplaces equipped with cameras (Wall Mounted, resp. Head Mounted Camera – WMC, resp. HMC) and virtual glasses that are also called Head Mounted Display (HMD). Such setup is necessary to immerse the user into augmented environment. The workspace is scanned by the camera mounted to the user's head and processed video sequence is projected back into HMD together with virtual object. Therefore, the user environment is extended by virtual objects which can be documents, construction models, schemas, pictures, etc. Other camera in front of the user scans whole workplace and creates user's virtual model, detects objects and tracks user hands. The workplace schema is depicted in Fig. 1:.

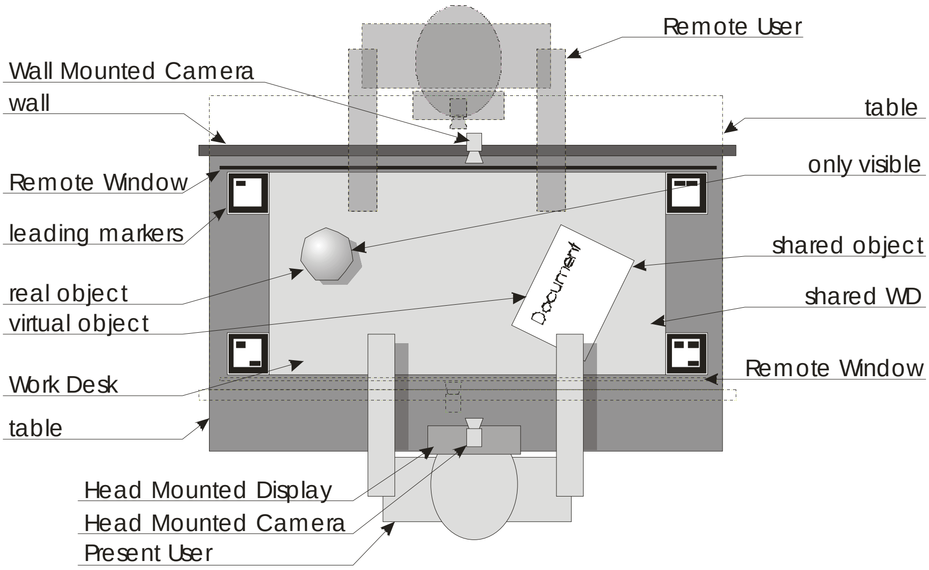

Fig. 1: *System workplace structure.*

At the very beginning, when the initiatory scene recognition is processed, the objects from the workplace are selected and labelled. The object class determines whether the object is going to be shared in cooperative environment and in what manner. The object structures, positions, textures and other parameters are stored in database. The engine working on database looks after the object properties consistency. Each user can modify the object properties, move the object over the workplace, change its content, etc. The object class determines whether and how the changes are shared between other users. The consistency engine also copes with concurrent object modifications that would lead to its inconsistency.

2.1 ENVIRONMENT SETUP

The step of environment initialization is crucial. The workplace is divided into Work-Desk and Remote Window. Work-Desk (WD) is determined to be shared with remote user and Remote Window (RW) shows the workplace of the cooperating user. In Fig. 1:., the special markers can be seen that help to detect WD and RW with sufficient accuracy and speed. The WD corner coordinates, respectively the relative corner positions, must be known.

Next critical step is setting of the lighting conditions. Hand, respectively finger shadows serve for hand pose detection, especially the distance above WD. Therefore, shadows must be detectable by the front camera. Hand tracking is then used for the basic virtual object manipulation. Further, the system detects real objects placed in WD and represents them as billboards. Then, the user can insert other virtual objects (documents, models …). All objects must be classified if they are supposed to be shared and shared mode must be selected (see Tab. 1: for more specification).

Real object	Virtual object
invisible object is not visible to other user	*invisible* …
only visible object is visible, but a remote user has no permission to manipulate the object in any manner	*only visible* …
shared copy the object is copied and remote user can manipulate it	*shared copy* remote user works with its own copy of the original object and is allowed to modify it in any manner (final synchronization with original object is question for future)
	shared object object is shared between both users with the same rights

Tab. 1: *Specification of object class modes.*

Each mode gives different rights to users to manipulate or modify the object. The most interesting mode, and unfortunately the most crucial one for consistency keeping, is 'Shared Object' when both users can cooperate and work with the same data.

2.2 SCENE RECOGNITION

Computer vision and image processing theories offer techniques for object detection, tracking and classification. When we were choosing which one we are going to use, the most important aspect, we took into the account, was the speed factor. Previous studies in virtual and augmented reality field showed that human, immersed in virtual, and especially in augmented environment, is much more sensitive for time delay than for space inaccuracy.

The background template is used as the reference image and all differences between the background template and actually processed image are held as potential objects. So far, we use simple static under-sampled background without any adaptation for global changes. Potential patches representing the differences are separately classified according to their size, color and shape. Objects are stored and tracked in the system, so the objects candidates from the actual image can be matched with those already existing in the system. It allows tracking of the objects and predicting their future positions for faster and more accurate detection in the next image.

The scene data are stored (using *Inventor scene graph* that is a toolkit for visualization and scientific simulation. During the setup of the augmented scene, the scene objects are loaded from VRML files. When no 3D object representation is available, only billboard representation is used. Then, objects are marked by their share mode and all necessary objects representations are exchanged between the computers.

The accuracy of the user and his hands detection is improved by the use of skin color model. Further, the known size of WD allows precise estimation of the object poses so they can be rendered in augmented scene in realistic way.

3 COLLABORATION

The design of the system and especially the shared object structure facilitates indispensable decrease of shared data amount. When the objects from WD are detected, they are distributed to the remote system with the background and the user model. Hereafter, only object poses and only changes within each object are distributed.

Sharing of the scene data is realized by the means of the replication. Each computer connected to the virtual environment can create replicas of the scene objects. And all computers connected to the virtual environment forms distributed system. Because of network latency and limited network bandwidth, dead-reckoning technique [3] is used. Network loading is reduced by prediction of the object movement by dead-reckoning algorithms. Moreover, frequency of updates may be lowered, and sending of the update may be delayed until predicted object position differs more then some threshold.

More complex problem arises, when more users want to modify the same object. Two different modifications on the same object on two different computers can break the scene consistency and may lead to unpredicted scene behavior. After studying how users are interacting with the objects, it can be said that only one user is usually interacting with the object. Therefore, ownership passing consistency is used for keeping the scene consistent.

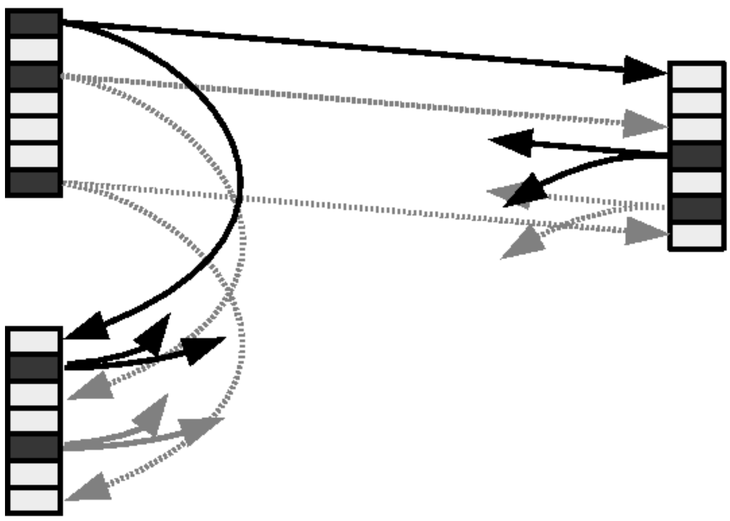

Fig. 2: *Consistency model.*

Ownership passing consistency means that each computer can have readable replica of the scene object. But there is only one writeable replica of the object throughout the whole distributed system. It means that if user wants to change something in the scene, it needs to gain ownership of the object. This means that the write permission is moved to its object replica. Such consistency may be very effective, because it does not require any network communication if the computer owns ownership of all required object replicas. Moreover in most situations, the object ownership is transferred just from time to time. So the performance of the application can be close to the standalone applications. Consistency model is showed on Fig. 2:, where darker items represents primary data and lighter ones read-only replicas. Arrows show the data flow direction of updating message.

When the scene object is changed, its new value is packed to the stream and sent to all the object replicas. The inherent Inventor scene saving and loading facilities are used, so the new value is stored in VRML format and transmitted. When exchanging the messages between objects was designed, another issue appeared – how to identify the objects throughout distributed system. It was decided to use Inventor functionality for naming nodes in the scene graph and each object can be identified by ip:name pair.

4 CONCLUSION

We have proposed the system for multi-user collaboration using augmented environment. Both system parts, scene recognition and augmentation and consistent data sharing, have several drawbacks and unsolved problems.

The object recognition is based on differences between background and actual image. Such approach is quite fast but the detection error, and especially misclassification rate, is not satisfactory. In worse environment conditions, such techniques can be almost unusable. Also problems with object-object occlusion and user-object can became unavoidable problem. Techniques such as interest point detectors or stable region detectors promise better results, but unfortunately with very high computational demands. Usage of specialized accelerating hardware for speed-up of scene recognition is promising.

REFERENCES

[1] Azuma, Ronald T.: A Survey of Augmented Reality, Teleoperators and Virtual Environments 6, 4 August 1997, pp. 355–385.

[2] Beran, V.: Augmented Multi-User Communication System, In: Proceedings of the working conference on Advanced visual interfaces, Gallipoli, IT, 2004, p. 257-260, ISBN 1-58113-867-9.

[3] Pečiva, J.: Collaborative Virtual Environments, In: Poster at MLMI'04 workshop, Martigny, CH, IDIAP, 2004, p. 1-1.

[4] Hesina, G., Schmalstieg, D., Fuhrmann, A., Purgathofer, W.: Distributed open inventor: A practical approach to distributed 3D graphics, Proceedings of the ACM symposium on Virtual reality software and technology, pages 74-81, http://www.cg.tuwien.ac.at/research/vr/div/vrst\\99.pdf, Mar. 1999. ACM Press