# The Mixed Reality Stage

# - an interactive collaborative pre-production environment


Wolfgang Broll
Fraunhofer FIT
Sankt Augustin, Germany
wolfgang.broll@fit.fhg.de

Wolfgang Krüger
Fraunhofer IITB
Karlsruhe, Germany
wolfgang.krueger@iitb.fhg.de

Stefan M. Grünvogel
Laboratory for Mixed Realities
Cologne, Germany
gruenvog@lmr.khm.de

Hans Krüger
Physikalisches Institut
Bonn, Germany
krueger@physik.uni-bonn.de

Martin Maercker
plan_b media ag
Cologne, Germany
maercker@planb-media.de

Frank Rohlfs
Theatertechnische Systeme
Syke, Germany
frank.rohlfs@ttssyke.de



**Abstract**

The Mixed Reality Stage is an interactive environment for supporting collaborative planning meetings in the application area of pre-productions. Based on augmented reality (AR) technology a group of experts shall be enabled to do complex planning and simulation tasks, without altering their current working habits radically. Virtual models and characters are projected into the real working environment of the users, enhancing the real world blending the real and the virtual. Multi-modal interaction techniques and tangible user interfaces are used to provide easy and natural access to virtual artifacts and characters.


## 1. Introduction

Creating a music show, a product launch on stage, a theatre or film production today requires a team of highly specialized professionals including directors, light and stage designers, actors and musicians. Often producers and customers are actively involved. Major costs arising from such a project are based on the time and the resources taken by the professionals to do their creative jobs. Here a good deal of discussions and costly false starts could be avoided if initial plans and ideas were visualized quickly, thus communicated more clearly, creating a common understanding and being evaluated by other participants in an early state.

In this paper we will first provide a general overview of the Mixed Reality Stage before we have a closer look into some of underlying concepts and technologies. In the third section we will report on our recent experience from user studies and discuss the results, before we conclude and have a look at future work in the fourth and final section.

**Figure 1** Real four-to-one model stage with two users

# 2. The Mixed Reality Stage

The Mixed Reality Stage is the sample application scenario used within the mqube (mobile multi-user mixed reality environment) project. It is based on a new innovative concept for supporting collaborative planning meetings using Augmented Reality (AR) technologies [1]. A real downscaled four-to-one model stage (cf. 1) is enhanced by virtual 3D objects such as stage props and characters representing actors or musicians. Professionals going to use the Mixed Reality Stage environment put on a personal display and wireless head-set (cf. 2) and can immediately start using the enhanced mixed reality environment. The personal displays projects virtual 3D objects into the real environment of the user providing the sensation of a fusion between the real and virtual environment.

**Figure 2** Personal display, tracking unit and head set

Details regarding the individual sub-components of the Mixed Reality Stage are described in the following sub-sections.

## 2.1. The Augmented Reality Framework

The Mixed Reality Stage application uses the AR framework Morgan – partially developed within the mqube project - as a general communication and visualization platform. As part of this framework, communication between all major system components is performed using CORBA. The framework also integrates the various input and manipulation components, including the user head tracking, camera-based object tracking, the light control, the stage control and the speech input. Other major components of the AR framework include the stereoscopic 3D rendering (viewer) and the API. The latter provides applications access to the virtual 3D objects as well as to the input devices connected to the framework.

The 3D visualization component renders the virtual scene components individually for each user's viewing position and orientation according to the received head tracking information. Optical see-through as well as video augmentation is currently supported. The latter one however, is used for demonstration purposes on external screens or projections only. In order to support proper occlusion between real and virtual objects, the visualization component provides support for phantom objects in both modes. As 3D scene description language VRML'97 is used. However, as the lighting support of this file format is not sufficient for describing more complex lighting conditions as required by a stage environment. Thus the existing standard as well as the export functions of the modeling tool 3d max has been extended accordingly. Additional extensions were made for easy bounding box visualization (used for highlighting) and the specification of phantom objects.

## 2.2. Light and Stage Control

The Mixed Reality Stage is directly connected to a professional lighting console and a stage control panel. The parallel real time controlling of show lighting and movable stage components in the Augmented Reality and on the model stage simulates and creates a real pre-production environment for the user – just like he knows from his daily work in 1:1 situations on stage.

The data exchange for the lighting control is based on the international standard DMX 512 signal for lighting. Currently the signal is transferred using a PCMCIA card with 1024 DMX channels. The system is based on standard DMX connectors, thus every professional lighting console can be plugged into the system just like the “real” world of show and entertainment lighting. Show controlling and light cues programming will be just as functional for the designers as on traditional 1:1 productions. Even when using the same console for the 1:1 production as on the Mixed Reality Stage, the board operator has the possibility to program the show beforehand. Thus by bringing its own controlling desk, a production can prepare the actual 1:1 show on their individual console. This is a major advantage of our approach, saving a lot of rehearsal time.

The electronic stage control is used for moving fly bars, point hoists, platforms, revolving stages, turntables and stage trolleys. The stage control includes one or several stationary or mobile operating consoles. Input devices are optionally touch screens, trackballs or joysticks. A PC keyboard is used for numerical and alphanumerical input. Alternatively, for numerical input, a special numeric dialog can be used via touch screen. Each monitor on the main console can be used for a different graphical view on the stage control condition. On a mobile console the user can switch between the different views. The mobile consoles are available with a 15” TFT touch screen, keyboard and two joysticks as well as a wireless console in size of a WebPad with two joysticks. The movement of the drives can be controlled by two joysticks or initiated by a START push button. The joysticks are used for controlling the direction and motor speed of the drives directly. The START button can be used to playback recorded movement of one or several drives between individual positions.  Several drives may be grouped and then controlled by one joystick or push button. One or several input computing devices are integrated in a console to evaluate the data entered by the user, like target positions, grouping of drives, movement mode, and performance data. The data is send to the Mixed Reality Stage.

## 2.3. User Head Tracking

The *BlueTrak* head tracking system [5] combines inertial sensor data and ultrasonic ranging measurements to achieve a six degrees of freedom information. The system is composed of two different types of modules: The *user modules*, and the *reference modules* (cf. 3). Both *user* and *reference modules* are based on a controller part, which consists of a microcontroller, a USB- and a Bluetooth-interface and an FPGA and analog circuits control the ultrasonic ranging measurements (cf. 4). Additionally, the user modules contain an inertial measurement unit, including accelerometers, gyroscopes and an electronic compass to determine its orientation. Both modules can be powered either by the USB connection or, if connected via Bluetooth, by a battery pack.

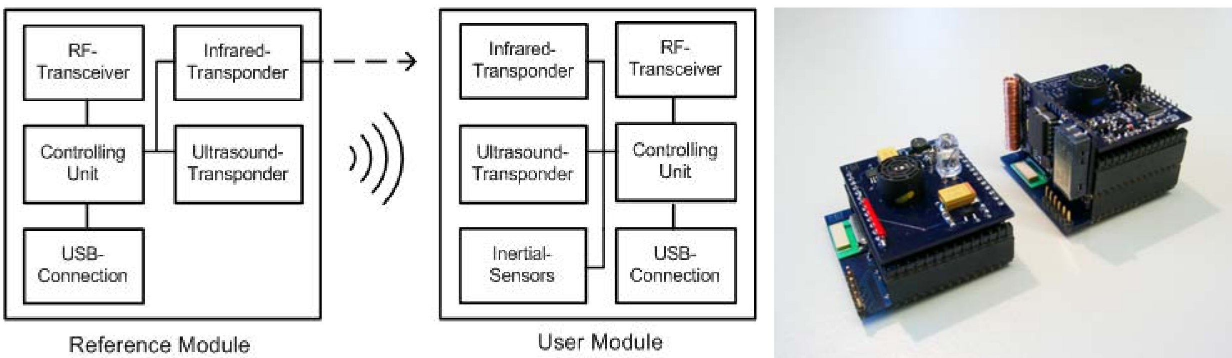


**Figure 3** Diagram and photo of the reference and user modules

The setup of the tracking system consists of a *user modules* fixed on each head-mounted display, four *reference modules* mounted on the traverses of the Mixed-Reality-Stage and a computer with USB or Bluetooth interface. The exact positions of the *reference modules* are determined during an easy calibration process. The references consecutively send ultrasonic pulses together with infrared pulses containing the ID of the reference. Measuring the time between the reception of infrared and ultrasonic pulses allows the *user module* to determine its distance to the sending *reference module*.

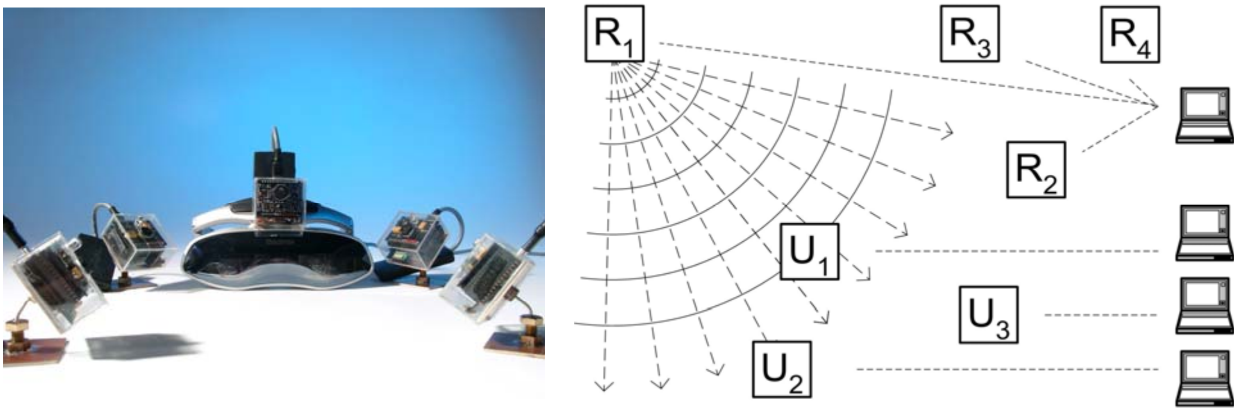


**Figure 4** Tracking system hardware (left) and system setup with every user module connected to its own computer (right); using this number of devices one computer would be sufficient

The number of reference modules is unlimited and allows an easy adjustment to the desired tracking volume and geometry (note that a user module does not need to be in the range of all reference modules at the same time). Due to this topology, the number of tracked devices is unlimited and has no impact on the update rate or the tracking accuracy. However, since the number of tracked devices connected to a single computer is limited, the system can be distributed on several machines (cf. 4). There can be different instances of the tracking program on separate computers. The program controlling the reference modules does not need to communicate with the user modules.

The measurements from the user modules are processed by an elaborate Kalman Filter. This Filter delivers accurate, robust values for the current position and orientation or for their estimated values at some future time (extrapolation). The orientation and position data of every user are transmitted via TCP/IP to the mqube augmented reality framework.

## 2.4. Computer Vision Based Object Tracking

Video-based object tracking is used to track in real-time hand-held objects (realoids) and to estimate their position and orientation (3D-pose with six degrees of freedom) in a world coordinate systems fixed to the stage. The estimated pose parameters are used for intuitive 3D-manipulation of associated virtual entities (virtualoids). Tracking and pose estimation rely on calibrated cameras and 3D-dimensional models of realoids. Thus monocular image sequences are sufficient. More complex and expensive multi-camera configurations with overlapping field of view (e.g. stereo) are not required. Instead multiple cameras can be used to enlarge the total field of view which has shown to be an important requirement.

Within the project IITB has developed and integrated flexible and easy to use tools for camera calibration, a marker-based object tracker for special realoids labeled by coded markers, and a more general edge-based object tracker which does not require marked realoids. All applications are based on standard PC-hardware and video cameras.

Internal camera parameters (e.g. focal length, lens distortion) are estimated from different views of a planar calibration target covered with coded markers. A similar planar calibration target is used for external camera calibration, i.e. to estimate the transformation from camera coordinates to stage coordinates.

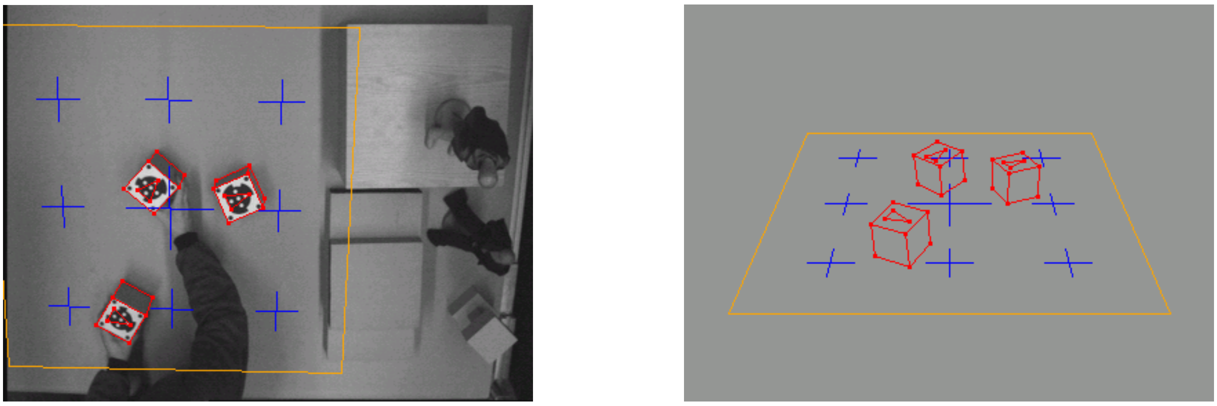

**Figure 5** Three tracked cubes labeled with different coded markers (left) and generated virtual front view (right). Overlays indicate the estimated pose (red) and the calibrated world coordinate system on the stage (yellow and blue).

The marker-based object tracker is able to estimate with sufficient accuracy the full 3D-pose from a single view of a planar marker which may cover only a small image area (e.g. a 25x25 pixel rectangle corresponding to a marker with 10cm diameter). Additionally the code is extracted and facilitates fully automatic initialization of object tracking as well as object identification. Different object trackers running on different computers and cameras may be combined to enlarge the total field of view. The trackers exploit bidirectional exchange of control information via TCP/IP to enable continuous object tracking across adjacent camera views. The current demonstrator has two marker-based object trackers each running on a 2 GHz Linux-PC. Each tracker is able to identify and track three marked cubes in real-time at 12.5 Hz (cf. 5).

**Figure 6** A tracked polyhedral object in front of background clutter (left) and a tracked free-form object (right). The red overlay indicates the estimated pose.

A different tracker which does not rely on coded markers estimates the 3D-pose of a realoid by maximizing the alignment between image edges and a projected 3D-model of that realoid. Video-based pose estimation for hand-held objects on a miniaturized stage is a challenging task due to a wide variety of illumination conditions, partial occlusions by hands and other objects, background clutter, and last but not least because objects tend to cover only a small image area. Considerable effort has been spend to achieve an acceptable degree of robustness under limited computational resources due to real-time requirements. The optimization process uses a multi-level random sampling strategy starting from straight line segments to generate promising pose hypotheses which are subsequently tested and refined on the level of extracted edge pixels. Figure 6 shows a tracked polyhedral object consisting of two cubes and a handle which is tracked in front of background clutter at a rate of 4.5 Hz on a 2 GHz PC. Trackable objects must generate sufficient edge structure to allow separation from background and reliable pose estimation.

To generate 3D-models for free-form objects a 3D-digitizer (Digiscan 2000) using structured light has been integrated. Figure 6 also shows an example for tracking a free-form object. The 3D-model for the dinosaur was created by the 3D-digitizer, automatically simplified and converted into the format used by the edge-based tracker.

Tracking has to be started from approximately known positions or initialized by user interaction (e.g. by using the view pointer). Development and integration of a robust edge-based object localization module would be an interesting topic for future work.

## 2.5. Ad hoc Character Animation

Creating character animation within an augmented reality system is a relatively new topic. The requirements in the context of the Mixed Reality Stage resemble scripting character animations sequences in games [3][4]. The user should be able to create, choose and delete different characters and there should be an easy way to create and edit animations, without low-level adjustment of the animations. To achieve this, the system provides the user with so-called subtasks, which model simple character behaviors like walking along a path, waving etc. To create a choreography (i.e. a complete animation sequence), the user moves the timestamp on a place on the timeline where he wants to execute a new subtask. After creating the new subtask with help of the menu, this new subtask is executed at the timestamp. Thus by moving forward and backward on the timeline and by placing commands on the timeline, the choreography of the character is created.

**Figure 7** Walking and waving Characters

**Figure 8** Character following a user-defined path (red) on the Mixed Reality Stage

The system also provides the user to create synchronization points, which are space-time constraints the character has to fulfill. If during the execution of a choreography the character detects, that he will not fulfill the space-time constraint (e.g. because the preceding choreography has changed), then the character interrupts his current tasks and jumps to the corresponding synchronization point. If on the other side, the character is too fast and has reached the synchronization point early, then the character waits until the specific time constraint is fulfilled.

The characters are implemented in a two layer architecture. The lower layer is responsible for the creation and manipulation of the motions in real-time (cf. 7). The higher layer represents the reflexive behavior exhibiting a fixed behavioral pattern in response to given stimuli (cf. 8).

## 2.6. Multi-modal User Interface

The interface consists of a set of generalized interaction mechanisms that were designed to be easy to learn and yet to allow, by incremental application, arbitrarily involved editing procedures [6]. Selection of virtual objects (*virtualoids*) is accomplished by utilizing head tracking data to determine which object is at the center of the user's field of view (*view pointer*). Operations on selected virtualoids are initiated by a small set of commands delivered by voice input.

The combination of view pointer and voice input leaves both the user's hands free to manipulate real objects (*realoids*) tracked by the object tracking component. The voice input "bind" creates an "interaction unit", a spatial association between a virtualoid and a realoid such that the virtualoid is kept in spatial sync with the realoid when the latter is moved [2].

Operations with a predominantly non-spatial domain (e.g. texture modification) are made available via a dynamic hierarchical menu associated with each virtualoid. Individual operations can also be represented as objects in the workspace. These "tool virtualoids" allow the user to

a) configure an operation's parameters once and then apply that operation to any number of virtualoids and

b) customize the workspace by loading and positioning tools as needed.

Tools can be made part of an interaction unit and their movement data interpreted as parameters to the operation.

**Figure 9** Selecting virtual menu entries using the view pointer

The implementation focus in the past year has been on multi-user interaction and on character animation mechanisms. The animation functionality provides for real time recording of animation paths using interaction units. The system translates the recorded path into a spline defined by control points which the user can edit individually e.g. by positioning them as part of an interaction unit. A "task bar" is instantiated for each character. It runs parallel to the system timeline and contains user-configurable actions that the character is to carry out at a certain time and at a certain position on the spline.

# 3. User Studies and Discussion

The user studies were conducted with seven accomplished professionals from the theatrical pre-production domain with an average professional experience of thirteen years. Their experience and integration in the social structure of the domain qualify them as representative of the individuals who decide on how a system like the mixed reality stage is to be integrated into the production process and under which conditions such a system is economically viable.

The sessions lasted two to three hours each and began with a demonstration of the system's basic functionality. The participants then worked with the system in the context of scenarios developed from their everyday practice. Individual features were discussed during this phase of the session. The participants were then asked to assess the strengths and weaknesses of the system as well as to compare it to other tools available to them.

A preliminary evaluation of the material gathered shows that the goals of the project were readily understood and seen as relevant to improving quality and efficiency in the domain. The individual interface mechanisms were generally well received. Some participants preferred the handheld track ball which was available as an alternative for navigating in meta-elements (e.g. menus) to the combination of view pointer and voice input.

The project's commercial success was seen as contingent on clients' acceptance of a pricing strategy based on higher personnel and facilities availability rather than based on the system's costs. Accordingly, participants cited an effective differentiation from low cost desktop 3D modeling applications as the greatest challenge. Since our approach could potentially have an impact on all of the processes central to the domain, an extended trial period (e.g. six months) under on-site conditions was seen as advisable for further qualification of the system's business prospects.

# 4. Conclusions and Future Work

In this paper we presented the Mixed Reality Stage – a multi-user environment to support collaborative planning tasks for pre-productions. We presented the underlying components and technologies required to create such an environment as well as the concepts used within the user interface. While limited to a rather small number of professionals, the user tests confirmed the validity of the initial approach and highlighted forthcoming directions and possible new application areas.

In our future work we will continue the evaluation with professionals from the pre-production domain, extending it into field trials at theatre locations. The multi-modal user interface will be further elaborated and its configuration facilities will be enhanced. Finally we plan to extend the overall project scope into new application areas, including but not limited to industrial plant planning, engineering applications and the preparation of exhibitions.

# 5. References


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

[2] Fitzmaurice, G., Ishii, H. Buxton, W. Bricks: Laying the Foundations for Graspable User Interfaces. Proceedings of the Conference on Human Factors in Computer Systems (CHI'95). p. 442-449. 1995.

[3] Grünvogel, Stefan M. and Schwichtenberg, Stephan. Scripting Choreographies, T. Rist et. al. (Eds.), IVA 2003, Lecture Notes in Artificial Intelligence 2792, pp 170-174, 2003, Springer-Verlag Berlin Heidelberg, 2003.

[4] Grünvogel, Stefan M. Dynamic character animations, International Journal of Intelligent Games & Simulation Vol.2 No.1 (2003), pp. 11- 19

[5] Krüger, H., Klingbeil, L., Kraft, E., and Hamburger, R. BlueTrak – A wireless six degree of freedom motion tracking system. In *Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, (Tokyo, Oct. 7-10, 2003), IEEE, Los Alamitos, p. 311-312.

[6] Lindt, I., Herbst, I., and Maercker, M. Interacting within the Mixed Reality Stage. In: Workshop Proceedings AVIR´03 / Magnenat-Thalmann, Nadia [Eds.], (2003).