Proceedings of 2004 IEEE/RSJ International Conference on
Intelligent Robots and Systems
September 28 - October 2, 2004, Sendai, Japan

Perspectives on Augmented Reality Based Human-Robot Interaction with Industrial Robots

Rainer Bischoff
KUKA Controls GmbH
Augsburg, Germany
RainerBischoff@kuka-controls.de

Arif Kazi
KUKA Roboter GmbH
Augsburg, Germany
ArifKazi@kuka-roboter.de

***Abstract*—First steps towards reliable Augmented Reality based human-robot interaction have been explored by the industrial robot manufacturer KUKA within the German cooperative research project MORPHA. Various aspects of Augmented Reality were analyzed and evaluated with respect to industrial requirements: interaction devices, tracking methods, accuracy, cost etc. As a result of this study, *training of, and qualification for, robot operation and programming* was selected as the most promising area for AR-based human-robot interaction with state-of-the-art AR techniques and devices. Therefore, research work concentrated on visualizing workflows that help inexperienced users to cope with rather complex robot operation and programming tasks. Several AR-based human-robot interaction prototypes were developed and presented to KUKA College students. Implementation details and results of initial experiments and a user survey are presented.**

Keywords - Augmented Reality; industrial robot operation and programming; human-robot interaction; training and education; user survey

I. INTRODUCTION

Augmented Reality (AR) is a growing and promising area in Virtual Reality (VR) research. An AR system generates a composite view for the user by combining the real scene viewed by the user and a virtual scene generated by a computer. Virtual information is embedded into the real world, thereby augmenting the real scene with additional information [Milgram & Kishino 1994].

While VR has entered product development long ago, AR is still an active area of research. In 1997, Azuma published a survey on AR [Azuma 1997] which was complemented in 2001 [Azuma et al. 2001]. In both publications the state of the art in AR components, open research questions and possible application scenarios in medicine, manufacturing, visualization, path planning, entertainment and military are described. Another complementary and excellent overview of current international activities can be obtained on the "Augmented Reality Page" [Vallino 2003]. To date, only a few applications of AR in the field of robotics or human-robot interaction have come to our knowledge. [Milgram et al. 1993] used AR in the context of robotic teleoperation to help a user understanding depth and other scene properties while controlling a robot. [Rastogi et al. 1996] augmented a wire frame model to the image of a real robot from a remote site. The wire frame model is used to simulate the robot movements at the local site before actually executing them at the remote site. [Giesler et al. 2004] use AR to interact with an autonomous mobile platform. Robot way points and route information can be given and overlaid to an operator's field of view by using voice and gestures.

Inspired by the AR enabling technologies and continuously increasing research activities in these fields, AR has also attracted KUKA's interest in its endeavor to further simplify human-robot interaction (cf. [Kazi & Seyfarth 2002], [Bischoff et al. 2002]). However, AR may not only be beneficial in creating human-friendly interfaces, but it holds promises throughout the life cycle of a robot (Figure 1). Starting with sales and planning, different robots could be virtually embedded into an existing manufacturing environment to check whether the working envelope and tools fit. Not yet produced tools could be virtually attached to a real robot flange to immediately start programming in the real environment. Robot installation could be simplified by displaying drill holes, mounting instructions and other information required for accurate robot placement. During set up workflows could be presented to guide personnel through the robot mastering process. Programming, training and operating of robots could be simplified by augmenting the position of coordinate systems, robot paths and other program information. Finally, service and maintenance are greatly simplified by presenting easy-to-follow instructions to service personnel.

Within the German cooperative research project "Anthropomorphic Robot Assistants" (MORPHA) KUKA

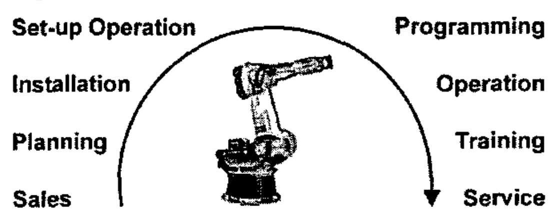

Figure 1: Augmented Reality can make life easier throughout the life-cycle of a robot system.

This work was partially funded by the German Federal Ministry of Education and Research (BMBF) under contract no. 01 IL 902 M 0 as part of the lead project MORPHA.

explored new ways of human-robot interaction [MORPHA consortium 2002]. The driving force behind this work is the fact that new markets and applications can only be conquered if robot programming and operation become easier. In these new markets, users are not expected to have fundamental robot know-how, and well-trained service personnel and programmers are not necessarily available. Within the fourth (and final) year of the MORPHA project KUKA first analyzed the requirements for using AR technology in manufacturing environments, then developed and built a first prototype, and eventually, carried out initial experiments to show the usefulness of AR in various human-robot interaction scenarios. KUKA closely cooperated with the ARVIKA consortium (another cooperative research project running in parallel to MORPHA) which concentrated on AR-based service and maintenance [ARVIKA consortium 2002].

In the following section basic requirements for an AR system are analyzed *with respect to an industrial robot manufacturer's perspective*. Section 3 describes the AR enabling technologies that have been chosen to develop an AR prototype, which could be used for various prototypical applications and first experiments (section 4). Section 5 reports on an advanced version of the AR interface that was installed in a KUKA robot training center. A user survey based on this interface was carried out to determine the suitability of AR-based human-robot interaction in general, and our approach in particular.

II. Basic Requirements for an AR System

Three major problems have to be solved in AR systems: (1) A user's pose in the world (to be more precise: his viewing position and direction) has to be identified and tracked, (2) the augmented image has to be computed (rendered) with respect to the current viewing pose and (3) an augmented visual stream has to be presented to the user. In this section, standard visualization methods and devices are presented first, followed by tracking methods and devices. All methods and devices are compared and evaluated with respect to their suitability for use in manufacturing environments. The ultimate goal of this comparison was to find a solution that provides sufficient flexibility for first experiments and can possibly be integrated in future intelligent teach pendants [Kazi et al. 2004].

A. Visualization methods and devices

Two main principles exist to augment the user's incoming visual stream: video-see-through (VST) and optical-see-through (OST) visualization. OST seems to be more appealing at first sight since the virtual information is merged through optical combiners (partly transmissive and reflective lenses) into the user's field of view. However, it has significant disadvantages: A believable augmentation is much harder to realize because it is very sensitive to calibration errors (matching errors of real and assumed looking direction, and size of the virtual image), and the augmentation computation has to account for head movements in real-time, i.e., has to predict the user's field of view based on past measurements.

These problems do not exist in video-see-through devices where the augmentation data can be perfectly synchronized with the video data. Furthermore, calibration is much easier done, especially if optical tracking methods are used, i.e., the camera that captures the real world is at the same time being used as a tracking device. Current VST devices are often virtual reality goggles (based on small LCD displays) that have been equipped with small video cameras. Still, there might be problems such as headache and motion sickness, or hand-eye coordination during fine manipulation, if the orientation of the cameras and the user's looking direction do not coincide.

In the course of study different visualization methods and devices have been evaluated by directly comparing products, product brochures and publications with respect to the following criteria (see Table 1):

- mobility range of a user wearing / carrying the device (e.g., is hands-free operation possible?);
- complexity of operation with respect to calibration procedures and overall handling;
- optical characteristics of classical representatives of commercially available devices;
- robustness and
- cost of such devices.

Since state-of-the-art Head Mounted Displays (HMD) developed for AR purposes are still fragile devices it can be assumed that they are not robust enough for, and would most likely not survive, harsh manufacturing environments. Furthermore, current HMDs have only poor optical characteristics: they possess a very limited field of view, a limited resolution (so far not exceeding SVGA), a limited number of colors (only one color with the retinal display) and insufficient brightness. The wearing comfort is also low because of their forward-oriented center of gravity and considerable weight when compared to normal glasses. Wearing HMDs at work can lead to muscle and skeleton pain, eye and vision problems and headache. Moreover, no standards are currently available that would allow an HMD manufacturer to certify its devices [Friedrich 2004], which is, however, an important aspect for future product development. Therefore, it was decided to opt for hand-

TABLE I. Comparison of Visualization Methods and Devices for Manufacturing Environments

visualization method	*mobility (hands-free)*	*operation (ease of use)*	*optics / perception*	*robustness*	*cost*
video-see-through HMD	++	o	o	o	o
optical-see-through HMD	++	-	o	-	-
retinal display (single color) HMD	++	-	-*	-	--
projector-based	–	++	+	++	o
handheld (monitor-) based	+(-)	++	++	+	+

*. Current displays limit the augmentation of realistic objects, but deliver sharp images.

held/monitor-based visualization, i.e., using the graphical display of the teach pendant or a separate screen for visualization purposes.

B. *Tracking methods and devices*

Accurately tracking the user's viewing position and orientation is crucial for a believable augmentation. An overview of recent tracking methods can be found, e.g., in [Rolland et al. 2001]. Similarly to our approach on comparing visualization methods and devices, we also compared different tracking methods and devices by evaluating products, product brochures and publications with respect to several criteria (Table II). These criteria are as follows:

- completeness of measurement data for pose information (ideally 3 DOF position and 3 DOF orientation);
- achievable accuracy of pose estimates (is largely determined by measurement principle);
- drift of sensors (e.g., pose estimates of inertial sensors drift without bound due to double integration of accelerations);
- robustness of tracking principle and devices with respect to environmental conditions in manufacturing cells (e.g., uncontrollable artificial magnetic fields);
- range (the mobility of the user is affected by the maximum allowed distance between tracking device and tracked feature);
- latency, i.e., elapsed time between start of measurement and availability of measurement data for the tracking algorithm;
- cost of commercial tracking devices.

The analysis of the basic requirements for an AR enabled robot interaction device made it clear that for the experiments, an optical tracking method would provide best accuracy and flexibility, while promising rapid development at reasonable cost, and still holding enough potential for future developments. Since marker-less and hybrid tracking methods are still in its early stages, it was decided to perform the optical tracking with the help of artificial markers that simplified the image processing and lowered the demands for computer power.

TABLE II. COMPARISON OF TRACKING METHODS AND DEVICES FOR MANUFACTURING ENVIRONMENTS

tracking method	*measurement data*	*accuracy*	*drift*	*robustness*	*range*	*latency*	*cost*
magnetic	+	+	-	-	o	+	o
acoustic	o	-	-	-	+	+	o
mechanic	+	+	+	+	-	+	+
optical	+	o	+	-	+	+	o
inertial	-	+/-	-	+	+	-	o
laser	o	+	+	-	o	+	-
GPS	-	-	-	+	+	-	+
hybrid	+	+	+	+	+	-	-

III. SYSTEM ARCHITECTURE AND IMPLEMENTATION

An AR system according to Figure 2 has been set up for first experiments. A video camera is used to capture real-world video data for visualization and computing camera pose information. This pose information is used to render the view of the virtual word, which is mixed with the video signal from the camera.

The decision in favor of marker-based tracking was supported by the fact that software packages were available for solving the camera pose estimation problem: "ARToolKit" developed at the University of Washington (http://www.hitl.washington.edu/artoolkit/) [Kato et al. 2000] and the "AR Browser" developed by Fraunhofer IGD within the German cooperative research project ARVIKA. After a short evaluation phase the AR Browser was selected. Although the ARToolKit was able to track arbitrarily shaped markers, the AR Browser provided simpler integration in custom programs and standard web browsers through Microsoft's ActiveX technology. Furthermore, a browser-based framework for implementing workflows is available, such as required for service or maintenance scenarios [Wohlgemuth & Triebfürst 2000].

Several two-dimensional markers (being neither axially nor rotationally symmetric) were constructed for tracking purposes. Each marker defines a single virtual world. One of the major problems to overcome was that 2-D markers as supported by the AR Browser proved insufficient to track randomly moving objects through space. Therefore, a three-dimensional cubic marker based on six standard two-dimensional markers was developed. It could reliably be tracked through space even when attached to a robot that moved arbitrarily.

Since three-dimensional markers were originally not supported by the AR Browser each 2-D marker had to be augmented by the same virtual world. Figure 3 shows different views onto the developed cubic marker augmented by an arrow and a plane covering each detected 2-D marker. This approach proved problematic when more than one side of the cube was visible at the same time due to calibration errors (i.e., mismatch between the real and the virtual world). Obviously, it would have been better to

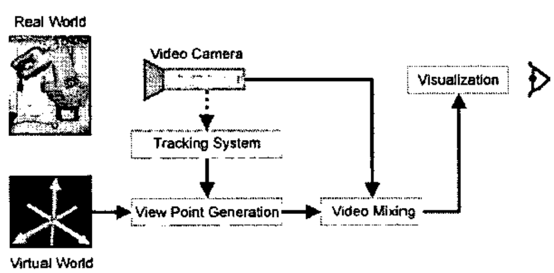

Figure 2: Working principle of a video-see-through or monitor-based AR system: A video camera is used to capture real-world video data for visualization. The video signal from the camera is mixed with rendered data from the virtual world model. The viewing direction needed to compute a believable view onto the virtual world objects is provided by a tracking system, which can be identical with the video camera used for obtaining real-world images.

Figure 3: Cubic marker made of six different planar markers. The virtual world of each marker consists of a colored 2-D square and an arrow perpendicular to a unique side of the cube. If the markers and the virtual and real worlds are not aligned properly a visible mismatch occurs.

define and to visualize only one virtual world according to the marker recognized best, or to fuse the available pose data appropriately, but none of these options was provided by the AR Browser. Another drawback was the lack of occlusion models which let the augmented objects appear unrealistic from some viewpoints.

A 1.6 MHz Pentium notebook with a video graphics card supporting OpenGL was used to run Microsoft's Internet Explorer Version 5.5 as a frame program for the AR Browser. Experiments with web cameras (USB connection) from three different manufacturers (Sony, Logitech and Philips) were carried out. Best results were achieved with CCD chip based cameras.

Preliminary experiments of running the AR software on a handheld low-cost WebPC were successful; however, the refresh rate was much lower than on the notebook. Main reasons were its rather low-performant CPU and a graphics card that did not support OpenGL. Consequently, the augmentation could only take place at about one image in 2-3s (instead of 10-20 fps achieved by the notebook) which proved insufficient for an intuitive operation.

IV. First Application Experiments and Results

Different applications scenarios throughout the before-mentioned life-cycle of a robot system (cf. Figure 1) have been evaluated in experiments and through in-house discussions. In the following, experiments and results in the areas of training, operation and programming, and service and maintenance are reported.

A. *Training of coordinate systems*

By attaching markers to the structural components of the robot it is possible to identify the six robot axes and display information about their axes of rotation and sense of orientation (see Figures 4a+b). This might help first time users of industrial robots to understand the kinematic structure of the robot. Besides training we strongly believe that any user who already knows the basics of robot operation could more easily and faster program the robot with the help of such augmented information. For instance, instead of trying out in which direction an axis moves (by pressing keys or other input devices in the manual jog mode), the augmented axis information could help a user to immediately rotate the axis in the wanted direction.

B. *Operation and Programming: Path information*

Attaching a three-dimensional marker to the robot could enable the visualization of current path information. The length of the arrow is meant to be an indication of the speed of the robot (Figures 4c+d). Furthermore, a path relative to a work piece could be overlaid. This could be useful, e.g., to help an operator to interactively teach the correct path to the robot (Figure 4e).

C. *Training of spacemouse usage*

Another interesting idea is teaching people how to use the teach pendant, in particular the spacemouse. The work-flow support by the AR Browser was used to implement an augmented spacemouse training scenario in which the user is asked to move the spacemouse in the indicated directions in order to move the robot accordingly (Figure 5).

D. *Service and maintenance*

One of the major topics in applied AR research is presenting information to the user that is helpful for service and maintenance. A Visual Diagnostic Tool (VDT) has been developed that is interfaced with the robot controller and guides the user in case of errors. Figure 6 depicts different visualization stages of the VDT after the emergency button on the KUKA Control Panel (KCP) has been pressed. In order to solve this "problem" several key presses are required (which even an inexperienced operator should know, but for simplification reasons we take this problem as an example). First, the user is informed about the problem by indicating which one of the robots needs attention. The "erroneous" robot turns red, and by clicking on it and pointing the camera towards the teach pendant, information on how to resolve the problem is displayed. On the right hand side a status message is displayed (here: "Emergency button has been pressed"), and at the bottom a message appears as to where the camera has to be pointed and what has to be done (here: "Point camera towards KCP. Turn emergency button in the indicated direction").

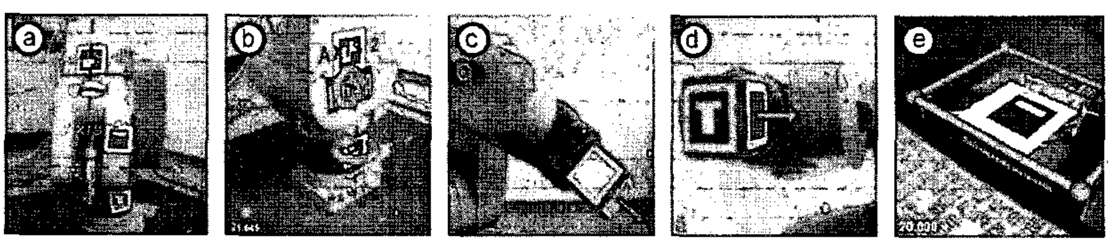

Figure 4: Different augmented visualization options on an industrial robot: (a, b) position of the axes and sense of rotation; (c, d) flange dependent visualization, e.g., current direction and speed of motion; (e) path with points to be taught on a workpiece.

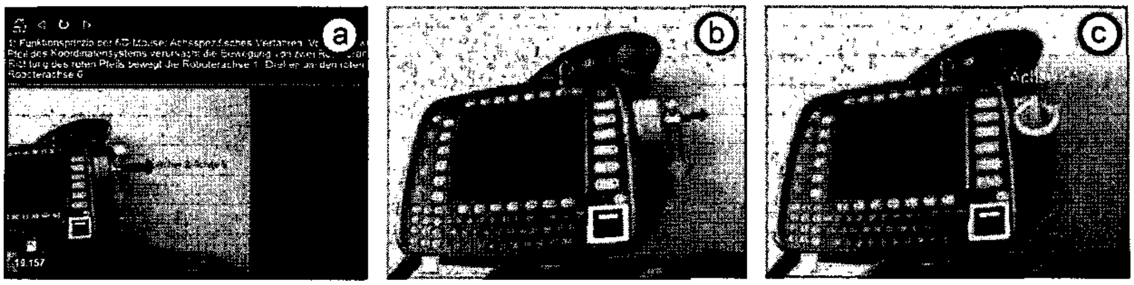

Figure 5: Learning how to use the spacemouse. (a) browser window with explanations and instructions; (b-c) arrows indicate the direction in which the user should push, pull or turn the spacemouse in order to perform a certain jog action.

After having turned the emergency button, the next screen is automatically displayed and instructions are given (here: "Switch to the message window (by pressing the button indicated by the left arrow), then acknowledge the reset of the emergency circuit (by pressing the button indicated by the right arrow)").

V. KUKA College Application and User Survey

A. *KUKA College application*

With respect to the life cycle of an industrial robot system and the obtained feedback from the initial experiments and use cases reported in section 4, the most promising area for exploiting AR information seems to be *training of personnel*. The main reasons are that the accuracy requirements are much lower than for all other fields of application, and even a single perspective or still pictures could provide a significant benefit for novice users, thus significantly lowering the demands on AR equipment.

The application developed for the KUKA training center ("KUKA College") is based on an advanced version of the AR Browser, but still following the system architecture depicted in Figure 2. As with the service and maintenance scenario described in the preceding section, the KUKA College application is running on a separate PC under Windows XP and is interfaced via (wireless) Ethernet to the robot controller for retrieving information about axis values and available (preprogrammed) coordinate systems. The application offers an easy-to-understand interface that allows students to select the coordinate system(s) to be displayed. At least one artificial marker has to be registered with the virtual model of the robot cell and aligned with the real cell to allow for a believable augmentation. To further enhance believability an occlusion model of the cell was provided to the AR renderer. Figure 7 shows a view of a standard KUKA training cell with augmented world, base and tool coordinate systems.

B. *User survey*

A user survey among KUKA College students was conducted to find out whether the augmentation helps them to understand the training material and the functioning of robot controller better. 23 students were interviewed of whom a large part (15 students = 65%) had never used industrial robots before, and the others had little working experience (1-2 years). The results are very encouraging. 78% of the students considered the augmentation of the robot's coordinate systems "important" or "very important"; nobody rated this feature as "unimportant". 74% of the students believed that AR could improve their understanding of the training material significantly; only 2 students considered AR supported education superfluous. Finally, a large number of 61% of the students believed that AR could also help them with their daily work with the robot.

VI. Summary and Conclusions

KUKA's research and development work in the German cooperative research project MORPHA was arranged around the development of a future generation of industrial human-robot interfaces. The driving force behind this work is that new markets and applications can only be conquered if robot programming and operation become easier. In these new markets, users are not expected to have fundamental robot know-how, and well-trained service personnel and programmers are not necessarily available.

First steps towards Augmented Reality based human-robot interaction for industrial robots were taken. The

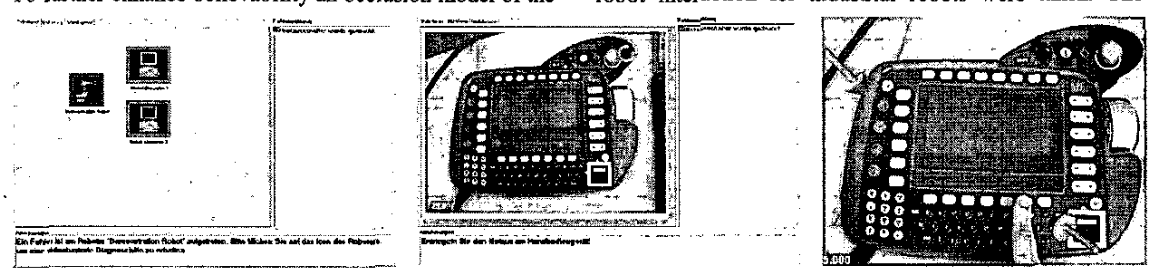

Figure 6: Visual Diagnostics Tool. (a) overview window indicates which robot has a problem and where to turn the camera; (b) emergency knob has to be turned counterclockwise to solve the problem; (c) certain key presses are required to resume operation.

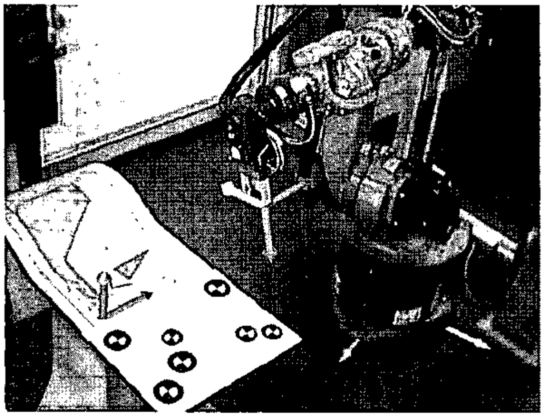

Figure 7: KUKA College demonstration. A standard KUKA training cell was equipped with an AR interface. Students were able to display world, tool and base coordinate systems, which enormously simplified their understanding of these coordinate systems and the related training material on robot operation and programming.

analysis of the basic requirements for an AR enabled robot interaction device made it clear that today's often used visualization devices for Augmented Reality, in particular Head Mounted Displays, are fragile devices with many cables and electronics that most likely would not long survive industrial working environments. Therefore, a handheld and monitor-based visualization and an optical tracking method were considered to provide best accuracy and flexibility, while promising rapid development at reasonable cost. It was decided that optical tracking was to be performed with the help of artificial markers that simplified the image processing and lowered the demands for computer power. The AR device used for the experiments was realized with a CCD based web camera that was attached to a standard notebook which was running a software package developed within the ARVIKA project.

Several AR interaction prototypes were developed that show the suitability of AR-based human-robot interaction, especially for relatively complex industrial robots. The prototypes concentrated on user training, programming and operation, and service and maintenance. Workflows were developed to guide novice users through control panel operation (emergency stop procedure, jogging the robot with the spacemouse).

It has been shown that AR has much potential for teaching students how to operate industrial robots. Although most of the experiments were targeted towards inexperienced users, a user survey in a KUKA robot training center and several in-house discussions revealed that AR holds much potential for a wide variety of industrial robot applications throughout the life-cycle of a robot from sales and planning to installation and programming to service and maintenance. It is expected that once the AR enabling technologies have matured AR can be robustly integrated in an intelligent teach pendant.

REFERENCES

ARVIKA Consortium (2002): Internet presentation of the BMBF lead project ARVIKA, available at www.arvika.de.

Azuma, R. (1997): A survey of augmented reality. Presence, Vol. 6, No. 4, pp. 355-386.

Azuma, R.; Baillot, Y.; Behringer, R.; Feiner, S.; Julier, S.; MacIntyre, B. (2001): Recent Advances in Augmented Reality. IEEE Computer Graphics and Applications Vol. 21, No. 6 (Nov/Dec 2001), pp. 34-47.

Bischoff, R.; Kazi, A.; Seyfarth, M. (2002): The MORPHA Style Guide for Icon-Based Programming. Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 02), Berlin, pp. 482-487.

Friedrich, W. (2004): ARVIKA – Augmented Reality für Entwicklung, Produktion und Service. Publicis Kommunikations-Agentur GmbH, Erlangen (in German).

Giesler, B.; Salb, T.; Steinhaus, P.; Dillmann, R. (2004): Using Augmented Reality to Interact with an Autonomous Mobile Platform. Proceedings of the International Conference on Robotics and Automation (ICRA 2004), New Orleans, to appear.

Kato, H.; Billinghurst, M.; Poupyrev, I.; Imamoto, K.; Tachibana, K. (2000): Virtual Object Manipulation on a Table-Top AR Environment. Proceedings of the International Symposium on Augmented Reality (ISAR 2000), pp. 111-119.

Kazi, A.; Seyfarth, M. (2002): Style Guide for Icon-Based Programming. Public-domain document in the BMBF lead project MORPHA. KUKA Roboter GmbH, Augsburg / Reis Robotics, Obernburg. Available at www.morpha.de.

Kazi, A.; Bunsendal, J.; Haag, D.; Baum, R.; Bischoff, R. (2004): Next generation teach pendants for industrial robots. In E. Prassler (ed.): Advances in Human-Robot Interaction. Springer Tracts in Advanced Robotics (STAR), Springer, Heidelberg, 2004 (in press).

Milgram, P.; Kishino, F. (1994): A Taxonomy of Mixed Reality Visual Displays, Transactions on Information Systems, Vol. E77-D, No. 12, pp. 1321-1329

Milgram, P.; Zhai, S.; Drascic, D.; Grodski, J. J. (1993): Applications of augmented reality for human-robot communication. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 1993), Yokohama, Japan, pp. 1467-1472.

MORPHA Consortium (2002): Internet presentation of the BMBF lead project MORPHA, available at www.morpha.de.

Rastogi, A.; Milgram, P.; Drasic, D. (1996): Telerobotic control with stereoscopic augmented reality. Proceedings SPIE: Stereoscopic Displays and Applications VII & Virtual Reality Systems III.

Rolland, J.P., Davis, L. Baillot, Y. (2001): A Survey of Tracking Technology for Virtual Environments. In: W. Barfield and T. Caudell (eds.): Fundamentals of Wearable Computers and Augmented Reality, Mahwah, NJ, 2001, pp. 67-112.

Vallino, J. (2003): Introduction to Augmented Reality. http://www.se.rit.edu/~jrv/research/ ar/introduction.html, last accessed November 25, 2003.

Wohlgemuth, W.; Triebfürst, G. (2000): ARVIKA: augmented reality for development, production and service. Proceedings International Conference on Designing Augmented Reality Environments (DARE 2000), Elsinore, Denmark, pp. 151-152.