A New AR Interaction Paradigm for Collaborative TeleAssistance system: The P.O.A

Sébastien Bottecchia, Jean-Marc Cieutat, Christophe Merlo, Jean Pierre Jessel

Noname manuscript No.
(will be inserted by the editor)

A New AR Interaction Paradigm for Collaborative TeleAssistance system: The P.O.A

Sebastien Bottecchia · Jean-Marc Cieutat · Christophe Merlo · Jean-Pierre Jessel

Received: date / Accepted: date

Abstract In this paper, we propose a prototype of a (collaborative) teleassistance system for mechanical repairs based on Augmented Reality (AR). This technology is generally used to implement specific assistance applications for users, which consist of providing all the information, known as augmentations, required to perform a task. For teletransmission applications, operators are equipped with a wearable computer and a technical support expert can accurately visualize what the operator sees thanks to the teletransmission of the corresponding video stream. Within the framework of remote communication, our aim is to foster collaboration, especially informal collaboration, between the operator and the expert in order to make teleassistance easier and more efficient. To do this we rely on classical repair technologies and on collaborative systems to introduce a new human-machine interaction: the Picking Outlining Adding interaction (POA interaction). With this new interaction paradigm, technical information is provided by directly Picking, Outlining and Adding information to an item in an operator's video stream.

Keywords Augmented Reality · Collaboration · Interaction · Teletraining · Teleassistance

S. Bottecchia · JM. Cieutat · C. Merlo
ESTIA, Technople Izarbel, 64210 Bidart, France
E-mail: s.bottecchia@estia.fr
E-mail: j.cieutat@estia.fr
E-mail: c.merlo@estia.fr

JP. Jessel
IRIT, Univ. Paul Sabatier, 31062 Toulouse, France
E-mail: jessel@irit.fr

1 Introduction

In this paper we focus on wearable computers as a relevant means of communication, especially with synchronous cooperation between stakeholders involved in remote repair operations through the use of distant video communication and augmented reality. In order to support such cooperation, we propose a new user-friendly human machine interaction: the Picking Outlining Adding (POA) interaction. The result is that the expert enhances the operators view in real time with augmentations (adding information) displayed on orthoscopic glasses that we have especially developed for this need.

In a worldwide context, companies must develop increasingly complex and innovative products in order to remain competitive. Nevertheless, a product which does its job alone is not sufficient and must innovate with regard to the services available, such as the maintenance process. In this context, having maintenance operators all other the world with the right level of knowledge and skills is clearly difficult and expensive but, with the increasing capacity of telecommunications and the explosion of the world wide web, teleassistance has become a significant challenge for industry. There are several business arguments in favor of teleassistance, like immediate real-time diagnosis without transport costs, or ever shorter repair and maintenance times (less waste of time and damage). Another argument is that complex mechanical repairs and assemblies require intensive operator training and teletraining is a suitable solution to reduce learning costs. Consequently, collaborative work between operators and experts presents a certain number of benefits, such as simultaneous quality control by the expert and manufacture feed-back by the operator.

Usually an expert provides oral explanations to a user on how to perform tasks, but for complex tasks, like for example, mechanical repairs or maintenance tasks, oral explanations are far from sufficient since the relevant technical information, human cooperation and guidelines are also rapidly required.

In complex industrial teleassistance, Augmented Reality (AR) techniques seem to be very promising and suitable with regard to supporting the requirements of new working methods able to enhance a user's perception by incorporating the computer-generated information (augmentations) into the real world. The first advantage is the possibility for an expert to see what the operator sees through a Head Mounted Display. In addition to usual voice communication, the other main advantage is the possibility for the expert to directly interact with the augmentations that the operator sees. Operators are thus provided with technical information and guidelines.

However, video link raises the problem of a synchronous communication. In addition, dealing with such multimodal information raises the problem of data merging to ensure consistency. The proposed new interaction paradigm aims to help us define solutions to these problems. We then have been implemented and experimented a new collaborative teleassistance tool based on AR.

2 From Maintenance to Collaborative Teleassistance

2.1 Well-known AR Systems for Maintenance Assistance

The first well-known AR prototype for maintenance assistance was KARMA [8] (Knowledge-based Augmented Reality for Maintenance Assistance). The tool was devoted to the maintenance of laser printers. The main features were the use of an Optical-See Through for display and ultrasonic sound sensors for tracking. In 1998 two AR prototypes for mechanical assistance were respectively developed by the Fraunhofer Institute and Boeing, the first one to assemble a bolt in a car door [16] and the second one to connect electric cables in a plane [13]. The common features of both prototypes was the use of fiducial markers for tracking. After 1998 ambitious AR programs, including industrial consortium and research laboratories, were started up, like ARVIKA [1], Starmate [20], nuclear power plant maintenance [7] and more recently AMRA [6]. Each program puts forward interesting points like a modular software framework for AR applications or indirect vision based on the use of tablet PC's. However, the most well-known AR systems for maintenance assistance do not propose real-time distant support collaboration facilities.

2.2 Collaboration Between an Expert and an Operator

Thanks to telecommunications, an operator can be assisted by a remote expert when performing difficult tasks. Cooperation is considered as an effective and concrete articulation between them, involved a collective action (in practice working towards a consensus). In fact, cooperation is an unpredictable and undetermined process [11] which influences the design and implemention of systems managing expert and operator interactions.

With regard to teleassistance, cooperation must be considered from the point of view of knowledge. The approach developed by [14] shows that knowledge management cannot be reduced to information management, but must consider the role of the stakeholders in the learning process. They are engaged in a process of successive tasks where both the expert and the operator have to interact to achieve them: at each stage they must perceive their immediate environment, analyse and decide what to do, and finally achieve it, following a repetitive cycle of perception, decision and action [9]. The expert is remote and must therefore be able to see what the operator sees in order to correctly perceive the situation. He must analyse the situation and make decisions based on both knowledge and experience. Although the expert's role is not to carry out repair tasks, he/she must transfer to the operator his understanding and the actions to be completed. The expert therefore needs powerful interactive tools to complement information transmitted by voice, such as multimodal addings managed in real-time. This enables formal tacit knowledge to be turned into explicit knowledge.

2.3 Original Work to Support Distant Collaboration

Normally there is formal knowledge with regard to industrial products: technical documents are available for operators as well as maintenance operations. Nevertheless, this is not enough to know how to use this knowledge. The expert helps the operator to transform this knowledge into skills by transferring tacit knowledge through adequate interactions.

When an expert is physically with an operator, he shows how to carry out tasks by doing them before the operator tries to do so (training through experience). In remote collaboration, the expert is unable to act alone, so the system to be implemented must provide similar

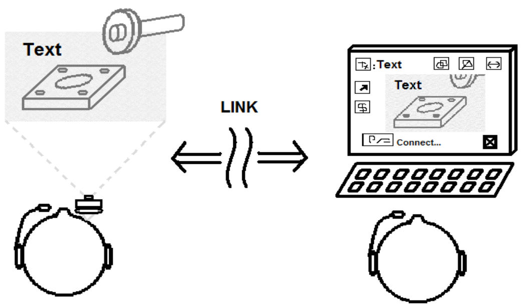

Fig. 1 POA interaction-based teleassistance enabled through Real-Time collaborative system.

interactions: picking an object, giving information on the object (what to do and why) and finally carrying out the action. The POA paradigm is based on these three steps:

– 'Picking' an object corresponds to selecting an object.
– 'Outlining' an object is a way to focus onto an object while giving oral explanations.
– 'Adding' information about actions illustrates what to be done by proposing 3D animations.

With a similar aim, [18] proposes using an active camera/laser pointer to show which object is to be taken or a direction to be followed. The relationship between Stakeholders is synchronous, but it is not yet possible to increase the user's sight. On the contrary, there are several systems, like [5] that enable an the expert to pick up an image in the operator's video stream, add augmentations to it and send back the result. Here the operator must wait for the static augmented image, but then the relationship between collaborators becomes asynchronous.

With our prototype we propose combining the advantages of classical repair systems and collaborative ones: the augmented field of view and synchronous relations thanks to the POA interaction.

Figure 1 describes the main principle of the system for teleassistance between an operator (left) and an expert (right): what is exactly seen by the operator's right eye is sent to the expert. The expert has a set of tools to send back formalized information in real time associated to what says and sees.

3 The proposed system

In this section, we shall describe the prototype built to support our POA interaction paradigm. The expert should be able to directly insert augmentation into the video stream to increase the operator's sight in real time and make him/her understand the tasks to be carried out.

3.1 Concept of the global architecture

In this context of remote assistance for mechanical repair, the ergonomics of augmented reality devices must provide a user-friendly interface for interlocutors. On the one hand, we need to define an architecture for transparent use by operators who must be freed from constraints related to the use of wearable computer technology equipment. On the other hand, the expert must be able to present information in a unique and intuitive way so as to grab the attention of the operator with efficiency. Directly clicking in the operator's video stream is a simple but efficient way to introduce guidelines. We therefore propose an architecture in figure 2 in order to support our POA interaction paradigm.

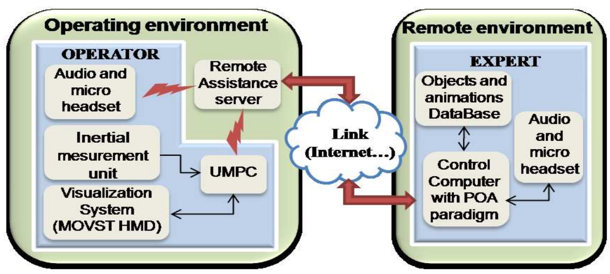

Fig. 2 Overview of the proposed architecture related to user's environment.

Thanks to the explosion of communication flows all over the world, we can use Internet as a connecting medium by which the operator receives the guidance of the expert. As communication possibilities are multimodal (visual, vocal and gestural), currently embedded operator systems are unable to process quickly enough to be fluid.

On the operator's site a remote assistance server will be in charge of carrying out complex calculations, managing the remote connection and enabling stakeholders to converse with an adequate system (i.e. wireless audio and microphone headset). As the operator should be the least instrumented, we opted for equipment consisting of a light visualization system (described in section 3.2) which will be coupled with an inertial measurement unit. This unit gives us the position and orientation of the operator's head at any given time so as to properly manage the virtual 'augmentations'. This equipment will be connected to a compact computer (like UMPC) able to perform light calculations. It also has a wireless link module in order to have an operating range with the remote server in the operating environment.

On the expert side, a simple computer is required to communicate with the operator. The computer will be able to establish a connection with the remote assistance server and will be equipped with software incorporating our POA interaction paradigm (described in section 3.3). Moreover, the computer manages a local database of the system's objects and animations to be maintained.

3.2 Visualization system (MOVST HMD)

Unlike virtual reality which requires total immersion, the operator using augmented reality will not lose perception of the real environment. This is why display systems in augmented reality must be mobile and not immersive. In order to achieve this, we often use devices like optical or video see-through head mounted displays (HMD), tablet PC's or more rarely projectors. Within the framework of teleassistance, the operator must be able to carry out work by apprehending the immediate real-world environment and keeping this hands free. That's why we have decided to use an HMD.

Therefore our goal is not to implement more intrusive HMD's for the operator since we are also trying not to overburden the operator's sight. Wrong perception in the increased visual area is potentially dangerous during a repair task. We therefore opted for a monocular HMD providing an 'augmented' window on the real world.

In order to maintain a "natural" visual field, the monocular HMD must make the operator think that he/she sees what would have been seen by the naked eye. Such systems are known as orthoscopic (Figure 3).

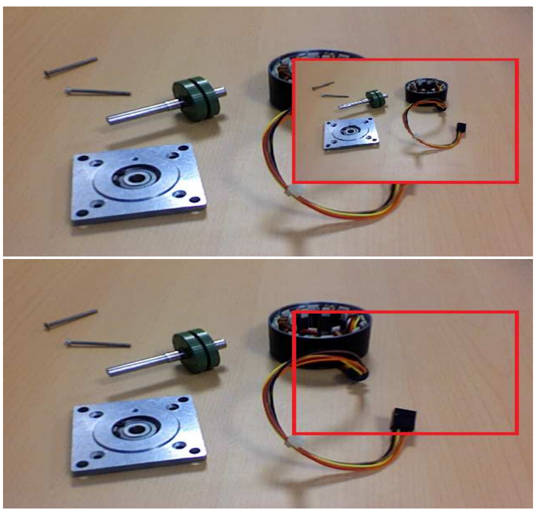

Fig. 3 Top: the human visual field with a classical view through a monocular HMD on the right eye. Bottom: visual field with an orthoscopic HMD (In red, the HMD's display limit).

Among the two existing HMD's for AR, Optical See-Through (OST) offers a direct view of the real world through a semi-transparent mirror. It also reflects images from a camera placed above it. Regarding the Video See-Through (VST), the real world is only seen through a camera. It therefore often raises a parallax problem because the camera is not always in the optical axis of the eye. With both systems, view acquisition of the camera usually leads to setting up a complex optical complex to produce [15,4]. Despite the fact that an OST provides a direct view of the real world, we chose the VST HMD. From our point of view, it has better display modes for our application. Here we can choose to use it as a system for increasing the user's sight as a classic display to present information not directly related to the point of view (explanatory video, technical plan). An OST is not so comfortable.

By integrating the fact that the device must be the least inconvenient and the lightest possible, we created our own system, a Monocular Orthoscopic Video See-Through HMD (MOVST HMD).

Our vision system is very simple to implement compared to existing ones because it does not use traditional optical devices to correct parallax, this function is computer corrected.

Our MOVST HMD (Figure 4) is entirely designed with commercial parts. The camera packaging used has been lightened to make it less cumbersome. The color camera sensor is a Sony HQ 1/3" Ex-View HAD PAL high resolution with a F2.0 and 30° field of view lens. Its total weight is under 30 grams. The display system is a ICUITI M920 weighing less than 100 grams, with a 24 bit color LCD VGA screen, 640x480 resolution and a 26° field of view. The camera is fitted directly on the display and is aligned with it to remove the parallax. We then created software that corrects the parallax with greater precision and which ensures the system's orthoscopy.

Fig. 4 Prototype of the Monocular Orthoscopic Video See-Through visualization system.

3.3 New modes of interacting

When we try to explain a predefined process to be followed to someone, an important aspect of the exchanged information consists of showing important parts of the conversation as visual objects or gestures. We generally use our finger to denote an object deserving attention [3], we outline a sketch of the scene to highlight important items and we even attempt to reproduce an animation of tasks to be completed. We propose implementing these principles by directly clicking on the operator's video stream (with the use of the Intel OpenCV library [10]). Thus, when experts "borrow" part of an operator's visual field through the MOVST HMD, they must be able to quickly "point out" what they want to show. This simple interaction can be modeled by an iconographic object (arrow, circle, etc). For the expert, by simply clicking with the mouse on the item of interest enables the operator to see the associated 'augmentation' (Figure 5) and quickly understand which object is being pointed out.

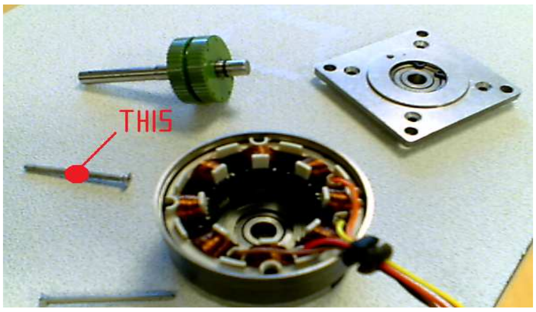

Fig. 5 'Augmented' view for the operator after 'Picking' by the expert.

This is the first type of interaction of the POA interaction paradigm: "Picking". The usual sentence "Take this object and..." becomes usable without ambiguity in a context of synchronous remote interaction. The operator can take an object and show it in front of the camera and ask, "What about this?".

With augmented reality, we have the possibility of going further than simple "Picking". We can add some interesting virtual items to the operator's field of view to underline vocal information given by the expert. However, visual realism is not always desirable in the scope of certain applications. A refined representation of information will not submerge the user with an excessive amount of detail. Thus, we can focus the user's attention on parts that deserve most interest, such as a highlighted object. In computer graphics, this type of rendering is called NPR (standing for Non-Photorealistic Rendering). Here we are mostly interested in highlighting objects by visualizing their shapes and outlines with sufficient visual information to catch the operator's attention.

This is the second type of interaction: "Outlining". Here the expert, with the "Picking" principle, points out an element whose 3D representation (modeled with CATIA and exported in VRML) is known by the system and contained into a database. The computer then recognizes the element and increases the sight of the operator by highlighting the silhouette or the outline of the highlighted object (red color in Figure 6). The objects involved in the concerned repair tasks are clearly identified during the explanation. It also helps to focus on the object to assemble or dismantle even if the connection conditions become difficult. (Object tracking and recognition use algorithms implemented in OpenCV)

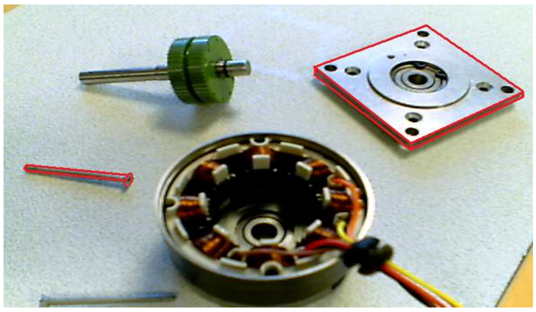

Fig. 6 'Augmented' view after 'Outlining' by the expert.

The "Outlining" arousing the interest of the operator on a particular object, the expert must then be able to mimic what to do. The use of animation seems therefore indicated.

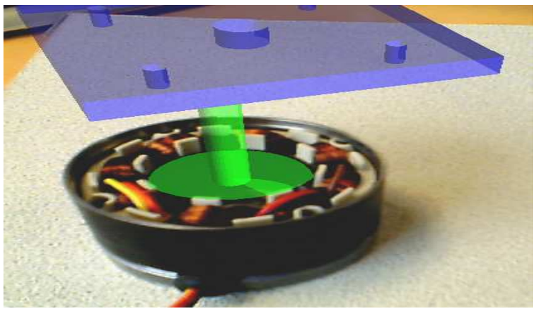

Fig. 7 'Augmented' view after an 'Adding' action by the expert.

Although the benefit of animations in understanding is not always demonstrated [12], the fact remains that its impact on explanation enhances the effectiveness of describing a process.

The third and final means of interaction is "Adding", which thus combines the "Outlining" principle with

2D/3D animation. The expert has a catalogue of animations (in VRML) directly related to the system subject to maintenance and to the objects being pointed out. He/she chooses the desired animation and then clicks on the video stream of the corresponding object. For example in Figure 7, the virtual items seen as outlines are animated to position them on the stator. (At the moment our implementation only searches for the oval representing the stator, calculating its position/orientation and matching it with the 3D animation).

4 Discussion

This initial experiment concerns a predefined assembly operation. The developed functions are hence limited to this situation. Although "Picking" is a simple operation that enables new and even unexpected situations to be taken into account, the two other POA paradigm concepts need to be enriched to do so. "Outlining" is not limited to visually identifying an object, but may be extended to show some of the object's properties, such as high temperature or fragile materials. Likewise, current "Adding" only supports the key step animations of the maintenance process. It can be extended to a catalogue of animated symbols to represent basic tasks and especially to combine them for unexpected tasks.

5 Conclusion and future work

In this paper we focus on synchronous and remote collaboration between an operator and an expert to complete maintenance and repair tasks. With our proposed AR system, we have introduced a new paradigm: "Picking Outlining Adding" interaction, in order to offer and implement an assistance environment that supports this collaboration. The presented prototype has been experimented on a predefined maintenance process and must be extended to support non-predefined repairs. To do so, the equipment must be improved. Furthermore, new interacting concepts and functions must be developed to embody informal knowledge exchanges.

References

1. Arvika, "augmented reality for development, production, servicing", http://www.arvika.de
2. R. Azuma, Recent Advances in Augmented Reality, IEEE Computer Graphics and Applications, 21, 34-37 (2001)
3. RA. Bolt, "Put-that-there": Voice and gesture at the graphics interface, SIGGRAPH '80: Proceedings of the 7th annual conference on Computer graphics and interactive techniques, 262-270 (1980)
4. O. Cakmakci and J. Rolland, Head-worn displays: a review, Journal of Display Technology, 2, 199-216 (2006)
5. P. Couedelo, Camka System, http://www.camka.com
6. JY. Didier and D. Roussel, AMRA: Augmented Reality assistance in train maintenance tasks, ISMAR'05: Workshop on Industrial Augmented Reality (2005)
7. AH. Dutoit and O. Creighton, Architectural Issues in Mobile Augmented Reality Systems: A Prototyping Case Study, APSEC '01: Proceedings of the Eighth Asia-Pacific on Software Engineering Conference, 341 (2001)
8. S. Feiner and B. Macintyre, Knowledge-based augmented reality, Commun. ACM, 36, 53-62 (1993)
9. JJ. Gibson, The theory of affordances, In "Perceiving, Acting and Knowing", Eds. R.E. Shaw and J. Bransford, Erlbaum (1977)
10. INTEL, OpenCV, http://sourceforge.net/projects/opencv/
11. J. Legardeur and C. Merlo, Empirical Studies in Engineering Design and Health Institutions, Methods and Tools for Cooperative and Integrated Design, 385-396, KLUWER Academic Publishers (2004)
12. RE. Mayer, When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction, Journal of experimental psychology, 11, 256-265 (2005)
13. U. Neumann and A. Majoros, "Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance", VRAIS '98: Proceedings of the Virtual Reality Annual International, 4 (1998)
14. I. Nonaka and H. Takeuchi , The Knowledge-Creating Company : How Japanese Companies Create the Dynamics of Innovation, Oxford University Press (1995)
15. N. Bretschneider, Head Mounted Displays for Fire Fighters, IFAWC, 3rd Inter. Forum on Applied Wearable Computing (2006)
16. D. Reiners and D. Stricker, Augmented reality for construction tasks: doorlock assembly, IWAR '98: Proceedings of the international workshop on Augmented reality : placing artificial objects in real scenes, 31-46 (1999)
17. J. Rolland and H. Fuchs, Optical versus video see-through head-mounted displays, chap.4 p113, Inc NetLibrary (2000)
18. N. Sakata and T. Kurata, Visual assist with a laser pointer and wearable display for remote collaboration, CollabTech06, 66-71 (2006)
19. B. Schwald, STARMATE: Using augmented reality technology for computer guided maintenance of complex mechanical elements, eBusiness and eWork Conference (2001)
20. B. Schwald and B. de Laval, An Augmented Reality System for Training and Assistance to Maintenance in the Industrial Context, 11th WSCG, vol.11, 1213 (2003)
21. A. State and KP. Keller and H. Fuchs, "Simulation-Based Design and Rapid Prototyping of a Parallax-Free, Orthoscopic Video See-Through Head-Mounted Display", ISMAR '05: Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality, 28-31 (2005)
22. D. Vinck, Intgration du savoir-faire. Capitalisation des connaissances, La connaissance: ses objets et ses institutions, 55-91, FOUET (1997)
23. X.W. Zhong and P. Boulanger, Collaborative Augmented Reality: A Prototype for Industrial Training, 21th Biennial Symposium on Communication (2002)