Author manuscript, published in "AH'2010 (Augmented Human), Megève : France (2010)"

hal-00585435, version 1 - 12 Apr 2011

T.A.C: Augmented Reality System for Collaborative Tele-Assistance in the Field of Maintenance through Internet.

Sébastien Bottecchia
ESTIA RECHERCHE - IRIT
Technopôle Izarbel
64210 Bidart (France)
(+33)5 59 43 85 11
s.bottecchia@estia.fr

Jean-Marc Cieutat
ESTIA RECHERCHE
Technopôle Izarbel
64210 Bidart (France)
(+33)5 59 43 84 75
j.cieutat@estia.fr

Jean-Pierre Jessel
IRIT
118, Route de Narbonne
31000 Toulouse (France)
(+33)5 61 55 63 11
jessel@irit.fr

ABSTRACT

In this paper we shall present the T.A.C. (Télé-Assistance-Collaborative) system whose aim is to combine remote collaboration and industrial maintenance. T.A.C. enables the copresence of parties within the framework of a supervised maintenance task to be remotely "simulated" thanks to augmented reality (AR) and audio-video communication. To support such cooperation, we propose a simple way of interacting through our O.A.P. paradigm and AR goggles specially developed for the occasion. The handling of 3D items to reproduce gestures and an additional knowledge management tool (e-portfolio, feedback, etc) also enables this solution to satisfy the new needs of industry.

Categories and Subject Descriptors

H.5.3 **[Information Interface and Presentation]**: Group and Organization Interfaces – *Synchronous interaction, Computer-supported cooperative work, Web-based interaction.*

K.4.3 **[Management of Computing and Information Systems]**: System Management – *Quality assurance.*

General Terms

Performance, Reliability, Experimentation, Human Factors.

Keywords

Augmented Reality – TeleAssistance – Collaboration – Computer Vision – Cognitive Psychology.

1. INTRODUCTION

Over the last few years the world of industry has held great expectations with regard to integrating new technological assistance tools using augmented reality. This need shows the difficulties encountered by maintenance technicians currently faced with a wide variety of increasingly complex mechanical/electronic systems and the increasingly rapid renewal of ranges.

The compression of training periods and the multiplication of maintenance procedures favor the appearance of new constraints linked to the activity of operators, eg. the a lack of "visibility" in the system to be maintained and the uncertainty of operations to be carried out. These constraints often mean that mechanics have to be trained "on the job ", which can in the long term involve a greater number of procedural errors and therefore increase maintenance costs as well as lead to a considerable loss of time.

In this highly competitive globalised context, the demand of industrialists to increase the performance of technical support and maintenance tasks requires the integration of new communication technologies. When an operator working alone needs help, it is not necessarily easy to find the right person with the required level of skill and knowledge. Thanks to the explosion of bandwidth and the World Wide Web, real time teleassistance is becoming accessible. This collaboration between an expert and an operator is beneficial in many ways, such as with regard to quality control and feedback, although a system enabling remote interactions to be supported is needed. With AR, we can now envisage a remote collaboration system enabling an expert to be virtually cop resent with the operator. By allowing the experts to see what the operators see, they are able to interact with operators in real time using an adequate interaction paradigm.

2. A.R. FOR MAINTENANCE & TELE-ASSISTANCE

We shall firstly take a brief look at existing systems and see that there are two major types which are quite separate. We shall then study the basic aspects which led us to build our solution.

2.1 Current systems

Amongst the AR systems aimed at assisting maintenance tasks, the KARMA prototype [8] is certainly the most well-known because it was at the origin of such a concept as far back as 1993. The aim of this tool was to guide operators when carrying out maintenance tasks on laser printers. Later other systems followed like those of the Fraunhofer Institute [20] and Boeing [18] in 1998. The purpose of the first was to teach workers specific gestures in order to correctly insert car door bolts. The second was aimed at assisting the assembly of electric wiring in planes. Following these systems, industry became increasingly interested in using such AR devices in their fields of activity. We then saw

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Augmented Human Conference, April 2–3, 2010, Megève, France.
Copyright © 2010 ACM 978-1-60558-825-4/10/04…$10.00.

hal-00585435, version 1 - 12 Apr 2011

the creation of more ambitious projects like ARVIKA [1] whose purpose was to introduce AR in the life cycle of industrial product, Starmate [22] to assist an operator during maintenance tasks on complex mechanical systems, and more recently ARMA [7] which aims to implement an AR mobile system in an industrial setting. Even more recently, Platonov [19] presented what can be described as a full functional AR system aimed at repairs in the car industry. This system stands out from others because it proposes an efficient technique enabling visual markers to be avoided.

The vocation of all of these systems is to support operators in the accomplishment of their tasks by providing contextualized (visual or sound) information in real time. Both of these conditions should reduce the risks of running errors according to Neumann's work [18].

Another common point is the importance placed on transparency in interaction with the machine. This is effectively a key point of AR in this field. Users must be able to pay their attention to the task in hand and not have to concentrate on how to use the tool itself, hence the different strategies of each project in creating prototypes. Also, the choice of the display device is important because the objective may be to reduce the need for resorting to classical supports (paper), thus leaving operator's hands free [24]. However, certain contradictory studies [25][10] are not conclusive with regard to the efficiency of AR compared to paper supports.

Finally, all these systems are particularly pertinent when tasks are governed by rules which allocate specific actions to specific situations, ie. within the framework of standard operational procedures. In this case we talk about explicit knowledge, although accessing this knowledge is not necessarily sufficient to know how to use it, which is known as tacit (or implicit) knowledge. This belongs to the field of experience, aptitude and know-how. This type of knowledge is personal and difficult to represent.

Thus, current AR systems for maintenance are of little use when an unforeseen situation occurs in which case it is sometimes necessary to resort to a remote person who has the required level of qualification.

It is only very recently that systems which support remote collaborative work for industrial maintenance have begun to appear. However, greater importance is given to the collaborative aspect than to maintenance. In [26] Zhong presents a prototype which enables operators, equipped with an AR display device to be able to "share" their view with a remote expert. The operator can handle virtual objects in order to be trained in a task which is supervised by an expert. However, the expert can only provide audio indications to guide the operator. Concerning [21], Sakata says that the expert should be able to remotely interact in the operator's physical space. This operator has a camera fitted with a laser pointer, and the entire system is motorized and remotely teleguided by the expert who can therefore see the operator's work space and point to objects of interest using the laser. The interaction here is therefore limited to being able to name objects (in addition to audio capabilities). There are other systems like [6] which enable the expert to give visual indications to an operator with an AR display device fitted with a camera. What the camera sees is sent to the expert who can "capture" an image from the video flow, add notes, then send back the enriched image to the operator's display device. Here the expert is able to enrich real images to ensure the operator fully understands the action to be carried out.

2.2 Motivation/Issue

In the paragraph above we saw that existing systems are either very maintenance-oriented with a single operator with a device or collaboration-oriented which do not necessarily enable direct assistance to be provided for the task in hand.

Our work is therefore based on the possibility of remote collaboration enabling both efficient and natural interaction as in a situation of copresence, whilst taking advantage of the possibilities offered by AR in the field of maintenance. Although in [14] Kraut shows us that a task can be carried out more efficiently when the expert is physically present, his study also shows that remote assistance provides better results than working alone, as confirmed by Siegel and Kraut in [23]. Other studies like [15] even show that a task can be accomplished more quickly and with less error when assisted rather than alone with a manual. However, communication mechanisms and the context play an important role when both the operator and expert share the aim:

- They share the same visual space. In remote collaboration, the expert does not necessarily have a spatial relation with objects [14] and must therefore be able to have a peripheral visual space so as to better apprehend the situation. This will directly affect coordination with the operator's actions and enable the expert to permanently know the status of work [9]. The lack of peripheral vision in remote collaboration therefore reduces the efficiency of communication when accomplishing a task [11].
- They have the possibility of using ostensive references, ie. deixis ("That one!", "There!") associated with gestures to name an object. Much research as in [14] and [4] suggests the importance of naming an object in collaborative work or not. This type of interaction is directly related to the notion of shared visual space referred to above.

These characteristics provided by a collaborative relationship of copresence are symmetrical [2], ie. those involved have the same possibilities. On the contrary, remote collaboration systems introduce asymmetries in communication. Billinghurst [3] highlights three main asymmetries which can hinder collaboration:

- implementation asymmetry: the physical properties of the material are not identical (eg. different resolutions in display modes)
- functional asymmetry: an imbalance in the functions (eg. one using video, the other not)
- social asymmetry: the ability of people to communicate is different (eg. only one person sees the face of the other)

Remote collaboration between an operator and an expert must be considered from the point of view of the role of each party, therefore necessarily introducing asymmetries, eg. due to the fact that the operator does not need to see what the expert sees. However, Legardeur [16] shows that the collaboration process is unforeseeable and undetermined, which means that experts may

hal-00585435, version 1 - 12 Apr 2011

have at their disposal possibilities for interaction close to those of operators as well as those which are available in real life, ie. the ability to name and mime actions. Finally, the underlying element with regard to collaboration in the field of tele-assistance is the notion of synchronism: collaboration may be synchronous or asynchronous. This shows the need for a real time interaction method between parties.

3. THE T.A.C. SYSTEM

3.1 Principle

To propose a solution combining remote collaboration and maintenance thanks to augmented reality, we have chosen two basic aspects:

- The mode of interaction between parties: This is the way expert can "simulate" their presence with operators
- The shared visual space: This is about being able to show the expert the operator's environment AND the way in which the operator is able to visualize the expert's information

Through these aspects we also suggest that our system is able to support synchronous collaboration between parties.

To implement this, we propose the following principle of use (figure 1): the operator is equipped with a specific AR display device. Its design enables it to capture a video flow of what the carrier's eye exactly sees (flow A) and a wide angle video flow (flow B). Amongst the two video flows which the expert will receive, there is the possibility of incrementing flow A thanks to our interaction paradigm (cf. paragraph 3.3). The incrementations are then sent in real time to the operator's AR display.

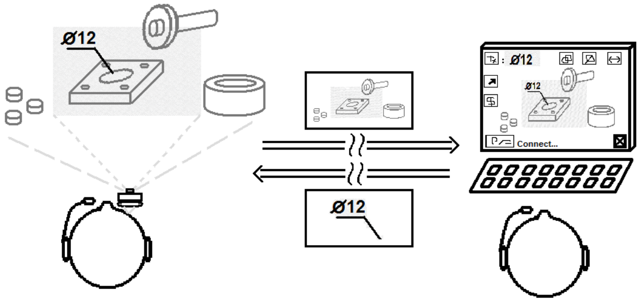


Figure 1. How the T.A.C. system works. The operator's view is sent to the expert who enhanced it in real time by simply clicking on it.

Hereafter we shall examine in greater detail our interaction paradigm and the visualization system supported by it as well as other functionalities.

3.2 Perceiving the environment

For each AR system developed, its type of display should be specifically chosen. Within the framework of maintenance, we must therefore take into account the constraints imposed by the operator's work. The many different aspects of using an AR system in working conditions linked to a manual activity poses certain problems. Furthermore, we must take into account how the situation is seen by the expert who must effectively apprehend the operator's environment as if he or she were there in person. In [5] we presented our visualization system carried by the operator and which is responsible for providing an exact vision of part of what is seen to the expert. This specific HMD, known as MOVST (Monocular Orthoscopic Video See-Through) satisfies the criteria of our application. The first of these criteria was that the operator must be able to easily apprehend the environment, without being immersed and keep as natural a field of vision as possible, ie. having the impression of seeing what can be seen with the naked eye (eg. orthoscopic).

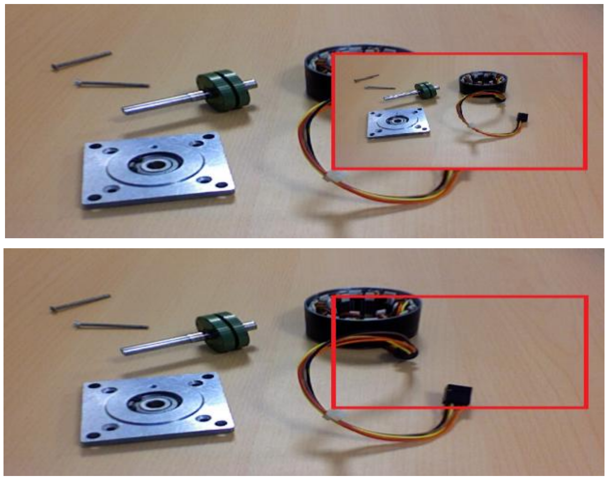
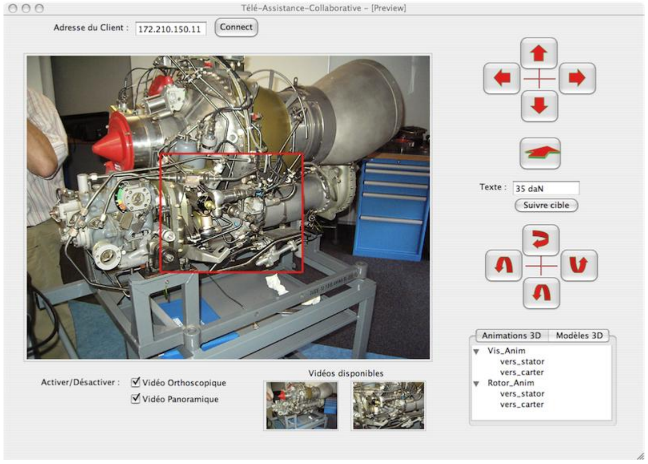


Figure 2. Simulation of the operator's field of vision carrying our MOVST. At the top a classic display (inside the red rectangle). At the bottom an orthoscopic display.

Figure 3. Prototype of our AR goggles known as MOVST.



Télé-Assistance-Collaborative - [Preview]

Adresse du Client : 172.210.150.11

Texte : 35 daN

Activer/Désactiver : ☑ Vidéo Orthoscopique ☑ Vidéo Panoramique

Figure 4. Expert interface. The orthoscopic view (inside the red rectangle) is placed in the panoramic view.

hal-00585435, version 1 - 12 Apr 2011

In order not to overload the operator's visual field with virtual elements, the choice of a monocular system has the advantage of only being able to be partly augmented. Finally, the "Video See-Through (VST)" principle was chosen for two reasons. Firstly, because it has an orthoscopic system, with a VST it is easier to implement the carrier's point of view. Secondly, it is possible to switch between orthoscopic display and classic display (figure 2). The advantages of the classic display lie in the fact that it can be used like any screen. It is therefore possible to present videos, technical plans, etc.

This so-called classic information is essential because it characterizes the "visibility" of the overall system subject to maintenance. Mayes in [17] distinguishes, amongst other things, the importance for the user of conceptualizing the task thanks to this type of information. However, the previous model of our MOVST only enabled the expert to see the "augmentable" part of the operator's field of vision, ie. approximately 30˚. In order to take into account the lack of peripheral vision as mentioned in 2.2, adding a second wide angle camera on the MOVST enables this problem to be solved (figure 3).

With regard to the expert's interface (figure 4), this gives a panoramic video of the scene in which the orthoscopic video is incrusted (PiP or Picture in Picture principle).

3.3 The P.O.A. interaction paradigm

In [5] we presented a new interaction paradigm based on the ability of a person to assist another in a task. Generally, when physically present together, the expert shows how to carry out the task before the operator in turn attempts to do so (learning through experience). To do this, the expert does not only provide this information orally as can be found in manuals, but uses more naturally ostensive references (since the expert and the operator are familiar with the context). Our P.O.A. (Picking Outlining Adding) paradigm is inspired by this and is based on three points:

- "Picking": the simplest way to name an object
- "Outlining": the way to maintain attention on the object of the discussion whilst being able to provide adequate information about it
- "Adding": or how to illustrate actions usually expressed using gestures

In order to implement these principles, we propose simply clicking on the video flow received from the operator.

The first mode, "Picking", therefore enables an element belonging to a work scene to be quickly named. This is equivalent to physically pointing to an object. The visual representation can be modelised in different ways like simple icons (circles, arrows, etc). Thus, the expert, by simply clicking on the mouse on an element of interest in the video, enables the operator to see the associated augmentation (figure 5). This provides experts with an efficient way of remotely simulating their physical presence in a more usual way and saying: "take this object and ...".

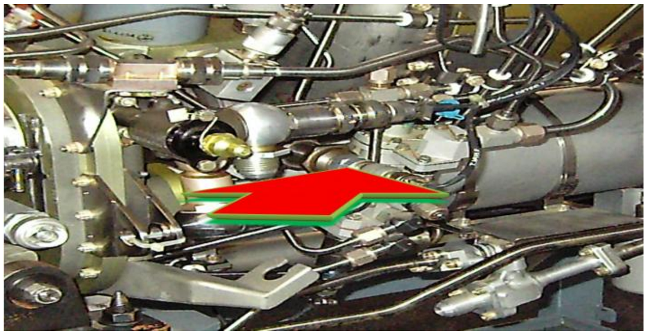


Figure 5. Operator's augmented view after a "Picking" operation. Here we clearly see the advantage of being able to discriminate an important element by showing it rather than describing it.

The second mode, known as "Outlining", uses the idea of sketching the elements of a scene using the hands to highlight them. These gestures support the verbal description. The principles of AR mean that we have the possibility of retranscribing this visually for the operator. Elements in the scene which require the operator's can be highlighted by drawing the contours or the silhouette of these objects (figure 6).

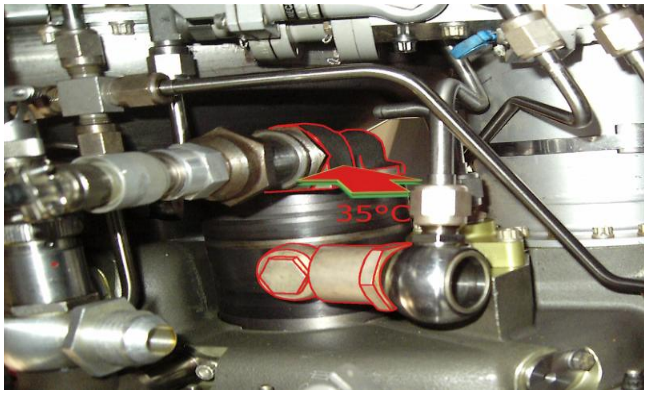


Figure 6. Operator's augmented view after "Outlining". The expert has selected the elements of interest and has given the temperature of an object.

With regard to the expert, this is done by clicking on the interesting parts whose 3D modeling is known by the system. We also have the possibility of adding characteristic notes (eg. temperature of a room, drill diameter).

The final mode, known as "Adding", replaces the miming of an action using adequate 3D animations. The expert has a catalogue of animations directly related to the system subject to maintenance. According to the state of progress of the task and the need, the expert can select the desired animation and point to the element to which it refers. Eg. (figure 7) the virtual element is placed directly where it should be.

hal-00585435, version 1 - 12 Apr 2011

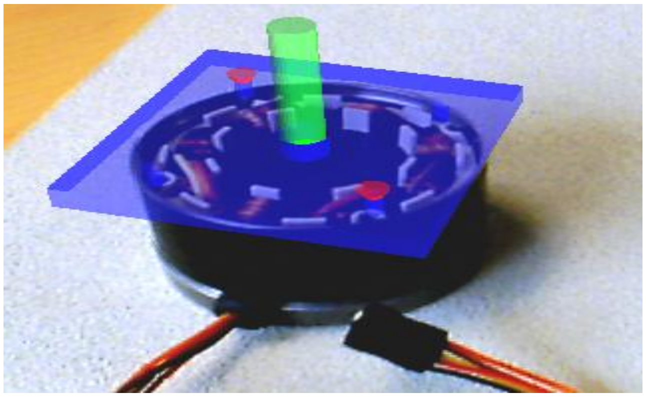


Figure 7. Operator’s augmented view after "Adding". The expert shows the final assembly using a 3D virtual animation placed on the real element.

3.4 Other functionalities

From the point of view of interaction by the system to support collaboration, P.O.A. interaction may be completed by the expert's ability to handle virtual elements. "Adding" enables actions expressed using gestures via animations to be illustrated, but this is only meaningful within the framework of a formal and therefore modelised process. This is not the case in unforeseen situations. For these, we are currently taking advantage of the formidable development of miniaturized inertial units. This works by handling this interactor associated with a 3D virtual element in the expert interface. The unit's position and orientation is retranscribed on the 3D element. The operator sees the virtual part handled just like if the expert had done so using the real part whilst using a tangible interface. However there is the problem of the expert not being able to handle at the same time both 3D interactors and the keyboard to provide important information. To support the transfer of implicit knowledge between the expert and operator, it is more efficient to add a "speech to text" type man-machine interaction mode.

The T.A.C. system, with its simulation of copresence, enables us to support a tool in full development in the world of work: the e-portfolio. This tool aims to manage a career path and validate acquisitions. In sum, this is a database enabling a person's skills to be capitalized. Thus, the T.A.C. system can be seen as a monitored system providing the possibility of recording images from different operations carried out with a view to an e-qualification. Work and qualifications can therefore be more easily combined.

Regarding the expert, recording images from different operations is first and foremost a quality control system. Since maintenance tasks in industry are highly formalized (set of basic operations), their supervision in the event of problems thanks to the synoptic view of operations carried out enables the cause to be analyzed. Its feedback can also be capitalized on to be used when designing future products and new maintenance procedures.

4. INITIAL RESULTS

4.1 Preliminary tests

We tested the T.A.C. system using two examples to verify their use within the framework of remote assistance. Operators do not have specific knowledge in the field of mechanical maintenance. The expert is someone who has been received a training in how to carry out maintenance on a helicopter turboshaft engine. The first example is not a real problem since it is simply question of assembling an electrically controlled engine in an order pre-defined by the expert (A, B, C, and D in figure 8). This simple example was initially chosen because 3D modeling and the associated animations were easy to create. Currently implementing our system is based on ARToolKit [13] and OpenCV [12] libraries for 3D recognition. To establish connection between two computers (voice and video session), we used the SIP signaling protocol implemented in SophiaSIP library. Transfer data is ensured by SDP and RTP protocols of the Live555 C++ library.

The second example concerns measuring the wear of blades in a helicopter turboshaft engine (E, F, and G in figure 9). This requires the use of a specific instrument which needs to be inserted in a precise location. The checking of measurements is supervised by the expert (this operation can prove delicate for beginners).

4.2 Discussion

During experiments, it became clear that our system provided easier and more natural interaction than other systems which provide traditional audio and video communication. The possibility for synchronous interactions by the expert vis-à-vis the operator stimulate exchanges and offer a strong feeling of being physically present which in the end leads to greater efficiency. This is due to the ability to act in unforeseen situations thanks to "Picking" and "Outlining" and well determined processes thanks to "Adding". Technical feasibility is extremely important with the increasing calculation capacities of laptops and the explosion of the bandwidth of communication networks. However, in experimental conditions the expert preferred it when the video offered a resolution of at least 640x480, which was not always possible because of our network's limited bandwidth. Most often, we were according the time of day forced to use a resolution of 320x240, enabling us to highlight this problem. It is therefore necessary to currently look at an exclusive communication solution between the expert and the operator. It also became clear that the expert would himself have liked to control the virtual objects supported by "Adding" instead of simple animations. We are currently working on this taking inspiration from interaction modes and virtual reality. Finally, the operator expressed the wish to be able to control switching from classic to orthoscopic displays in the MOVST and more generally have greater possibilities for controlling the display system.

5. CONCLUSION

In this paper we have presented a system enabling two remote parties to be able to collaborate in real time in order to successfully carry out a mechanical maintenance task. This system is based on our P.O.A. interaction paradigm enabling the expert's presence to be simulated with an operator in a situation of assistance. This prototype was tested on simple cases, but which were representative of certain real maintenance tasks and it showed that it was able to support both defined and undefined interaction processes. However, we must provide the means for greater interaction between parties and carry out a more in-depth study of the real benefits of such a system.

hal-00585435, version 1 - 12 Apr 2011

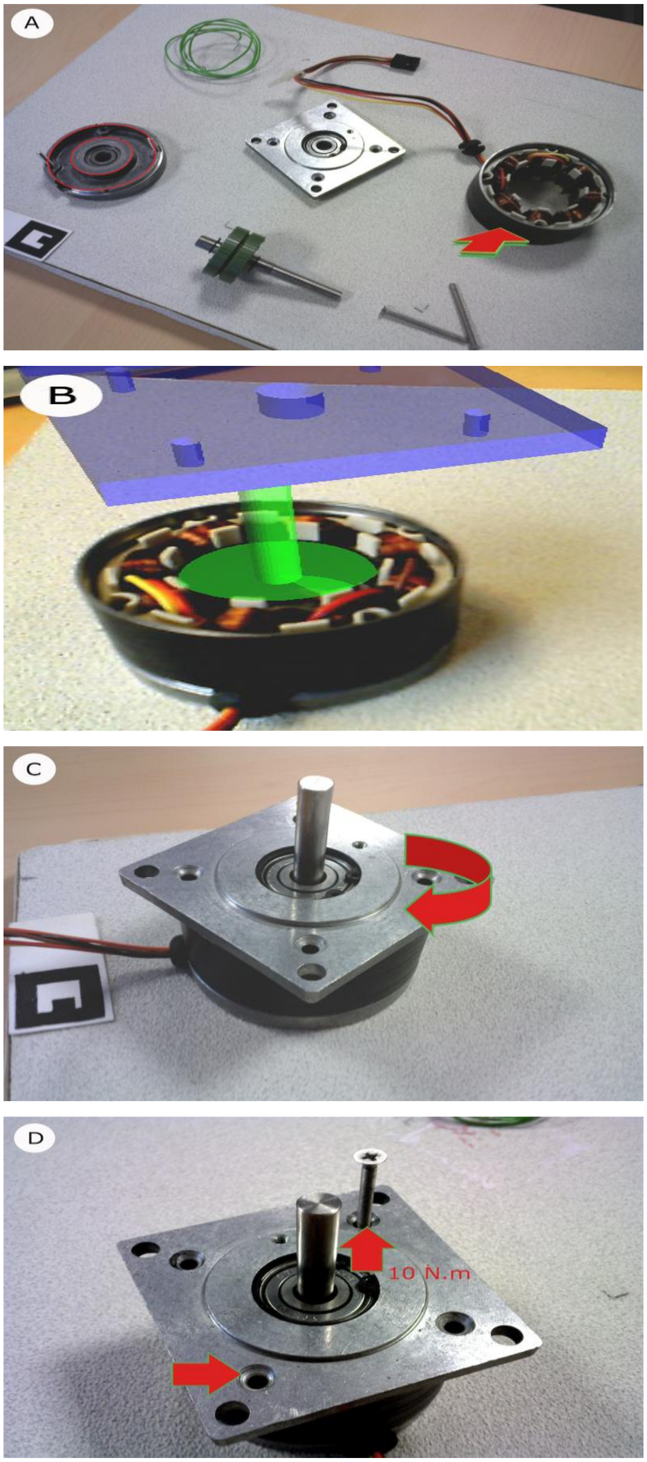


Figure 8. Example of collaboration

A: "Take this stator and put it on the red support"

B: "That's how the rotor and the case are put together"

C: "Turn the carter in this direction until you hear it click"

D: "Put the screws here and here with this torque"

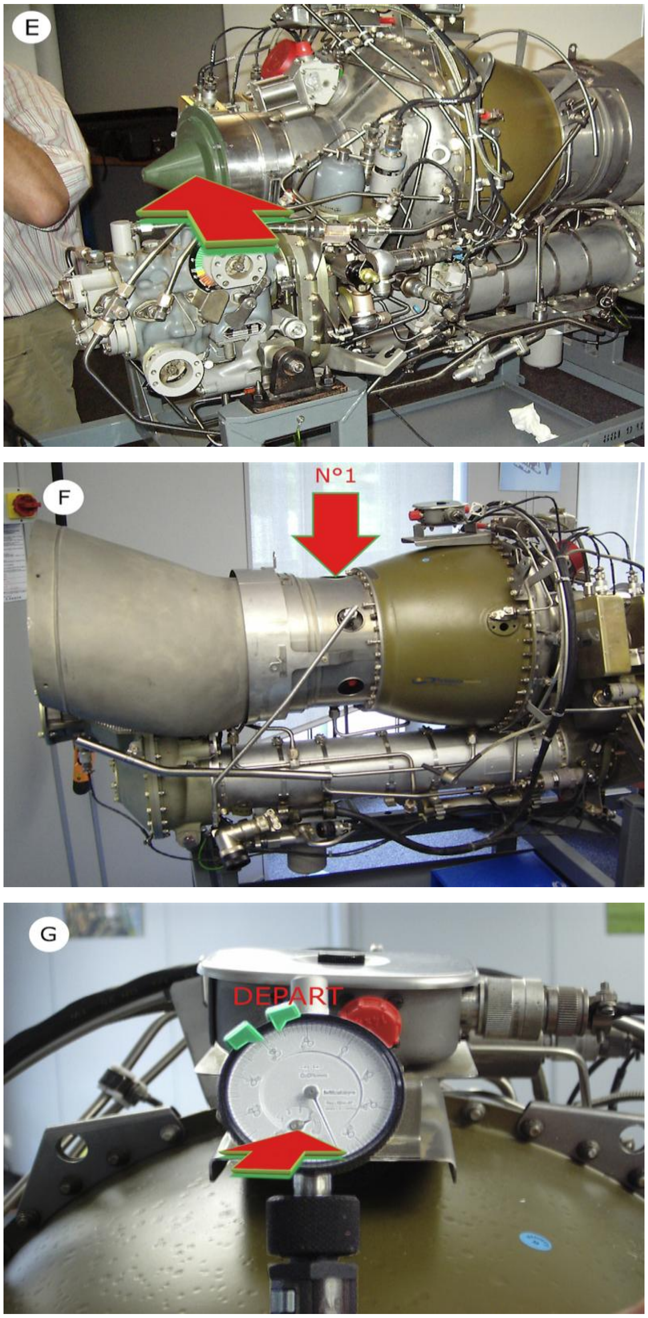


Figure 9. Other examples of assistance.

E: "Undo this cap so you can then turn the shaft"

F: "Place the instrument in hole no. 1, that one there"

G: "Look over here, the small needle says 2 tenths, that's ok"

6. ACKNOWLEDGMENTS

We would like to thank LCI, which specializes in turboshaft engine maintenance, Christophe Merlo for his help on the knowledge of collaborative processes and Olivier Zéphir for his various contributions with regard to cognitive psychology.

hal-00585435, version 1 - 12 Apr 2011

7. REFERENCES

[1] Arvika. Augmented reality for development, production, servicing. http://www.arvika.de, URL.

[2] Bauer, M., Heiber, T., Kortuem, G. and Segall, Z. 1998. A collaborative wearable system with remote sensing. ISWC '98: Proceedings of the 2nd IEEE International Symposium on Wearable Computers, page 10.

[3] Billinghurst, M., Kato, H., Bee, S. and Bowskill, J. 1999. Asymmetries in collaborative wearable interfaces. ISWC '99, pages 133–140, 1999.

[4] Bolt, R.1980. 'put-that-there': Voice and gesture at the graphics interface. SIGGRAPH '80: Proceedings of the 7th annual conference on Computer graphics and interactive techniques, pages 262–270.

[5] Bottecchia, S., Cieutat, J., and Merlo, C. 2008. A new AR interaction paradigm for collaborative TeleAssistance system: The P.O.A. International Journal on Interactive Design and Manufacturing, N°2.

[6] Couedelo, P. Camka system. http://www.camka.com, URL.

[7] Didier, J. and Roussel, D. 2005. Amra: Augmented reality assistance in train maintenance tasks. Workshop on Industrial Augmented Reality (ISMAR'05).

[8] Feiner, S., Macintyre, B. and Seligmann, D. 1993. Knowledge-based augmented reality. Commun. ACM, 36(7):53–62.

[9] Fussell, S., Setlock, L., Setlock, L.D. and Kraut, R. 2003. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 513–520.

[10] Haniff, D. and Baber, C. 2003. User evaluation of augmented reality systems. IV '03: Proceedings of the Seventh International Conference on Information Visualization, page 505.

[11] Heath, C. and Luff, P. 1991. Disembodied conduct: Communication through video in a multi-media office environment. CHI 91: Human Factors in Computing Systems Conference, pages 99–103.

[12] INTEL. OpenCV. http://sourceforge.net/projects/opencv/

[13] Kato, H. and Billinghurst, M. ARToolkit. http://www.hitl.washington.edu/artoolkit/, URL.

[14] Kraut, P., Fussell, S. and Siegel, J. 2003. Visual information as a conversational resource in collaborative physical tasks. Human-Computer Interaction, 18:13–49.

[15] Kraut, R., Millerand, M. and Siegel, J. 1996. Collaboration in performance of physical tasks: effects on outcomes and communication. CSCW '96: Proceedings of the 1996 ACM conference on Computer supported cooperative work, pages 57–66.

[16] Legardeur, J. and Merlo, C. 2004. Empirical Studies in Engineering Design and Health Institutions, chapter Methods and Tools for Co-operative and Integrated Design, pages pp. 385–396. KLUWER Academic Publishers.

[17] Mayes, J. and Fowler, C. 1999. Learning technology and usability: a framework for understanding courseware. Interacting with Computers, 11:485–497.

[18] Neumann, U. and Majoros, A. 1998. Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. VRAIS '98: Proceedings of the Virtual Reality Annual International Symposium, page 4.

[19] Platonov, J., Heibel, H., Meyer, P. and Grollmann, B. 2006. A mobile markless AR system for maintenance and repair. Mixed and Augmented Reality (ISMAR'06), pages 105–108.

[20] Reiners, D., Stricker, D., Klinker, G. and Muller, S. 1999. Augmented reality for construction tasks: doorlock assembly. IWAR '98: Proceedings of the international workshop on Augmented reality: placing artificial objects in real scenes, pages 31–46.

[21] Sakata, N., Kurata, T., Kato, T., Kourogi, M. and Kuzuoka, H. 2006. Visual assist with a laser pointer and wearable display for remote collaboration. CollabTech06, pages 66–71.

[22] Schwald, B. 2001. Starmate: Using augmented reality technology for computer guided maintenance of complex mechanical elements. eBusiness and eWork Conference (e2001), Venice.

[23] Siegel, J., Kraut, R., John, B.E. and Carley, K.M. 1995. An empirical study of collaborative wearable computer systems. CHI '95: Conference companion on Human factors in computing systems, pages 312–313.

[24] Ward, K. and Novick, D. 2003. Hands-free documentation. SIGDOC '03: Proceedings of the 21st annual international conference on Documentation, pages 147–154.

[25] Wiedenmaier, S. and Oehme, O. 2003. Augmented reality for assembly processes-design an experimental evaluation. International Journal of Human-Computer Interaction, 16:497–514.

[26] Zhong, X. and Boulanger, P. 2002. Collaborative augmented reality: A prototype for industrial training. 21th Biennial Symposium on Communication, Canada.