Consistency in Augmented Reality Systems

Emmanuel Dubois[1,2], Laurence Nigay[1], Jocelyne Troccaz[2]

[1] CLIPS-IMAG, IIHM, 385 rue de la bibliothèque, BP 53,
F-38041 Grenoble Cedex 9, France
[2] TIMC-IMAG, GMCAO, Faculty of Medicine (IAB), Domaine de la Merci,
F-38706 La Tronche Cedex, France
{Emmanuel.Dubois, Laurence.Nigay, Jocelyne.Troccaz}@imag.fr

Abstract. Systems combining the real and the virtual are becoming more and more prevalent. The Augmented Reality (AR) paradigm illustrates this trend. In comparison with traditional interactive systems, such AR systems involve real entities and virtual ones. And the duality of the two types of entities involved in the interaction has to be studied during the design. We therefore present the ASUR notation: The ASUR description of a system adopts a task-centered point of view and highlights the links between the real world and the virtual world. Based on the characteristics of the ASUR components and relations, predictive usability analysis can be performed by considering the ergonomic property of consistency. We illustrate this analysis on the redesign of a computer assisted surgical application, CASPER.

1. Introduction

Integrating virtual information and action in the real world of the user, is becoming a crucial challenge for the designers of interactive systems. The Augmented Reality (AR) paradigm illustrates this trend. The main goal of AR is to add computational capabilities to real objects involved in the interaction. The *Augmented Paper Strip* [15] is an example of such an attempt in the air traffic control domain: the goal is to add computational capabilities to traditional paper strips. Another system called *KARMA* [11] provides information to a user repairing a laser printer, by indicating with 3D graphics which parts of the printer to act on, according to the defined maintenance process. The *Tangible Interface* [14] constitutes another example of the combination of real and virtual entities: everyday life objects are used by the user to interact with the computer. More examples of AR systems are presented in [2]. In [9], we have already illustrated this wide area of interactive systems and highlighted an important classification characteristic:

- some systems (*Augmented Reality systems, AR*), enhance interaction between the user and his/her real environment, by providing additional computer capabilities or data,
- while others (*Augmented Virtuality systems, AV*) make use of real objects to enhance the user's interaction with a computer.

Since our application domain is Computer Assisted Surgery (interaction with the real world, the patient, enhanced by the computer), we particularly focus on the characterization and description of *Augmented Reality systems*. One of the main design challenges of such *Augmented Reality Systems* (*AR*) is to merge real and virtual entities. Real environment and real entities are prerequisite for the design of such systems. The composition of these two kinds of entities constitutes the originality of *AR* systems. In addition, information or action are defined by the *AR* system to facilitate or to enrich the natural way the user would interact with the real environment. Consequently, the main point of interest during the design should be the outputs of the systems, so that additional information and action are smoothly integrated with the real environment of the user.

In this paper, we first briefly present our application domain: Computer Assisted Surgery. Through the presentation of the taxonomy of the domain, we motivate our approach that aims at studying the two types of entities (real and virtual ones) involved in the output user interfaces of *AR* systems. We then present our notation called ASUR. ASUR is based on the principles of the notation OP-a-S [9], which is enriched by characteristics that describe the user's interaction. The ASUR notation describes a system with a user's task-centered point of view and highlights the links between the real world and the virtual world. Based on our descriptive notation, we then show how predictive evaluation of the consistency ergonomic property can be addressed. We illustrate our approach through the redesign of *CASPER*, a computer assisted surgery system.

2. Motivation and Application Domain

There are many application domains of *Augmented Reality*, including construction, architecture [20] and surgery [3], [6], [18]. Our application domain is *Computer Assisted Surgery* (*CAS*). The main objective of *CAS* systems is to help a surgeon in defining and executing an optimal surgical strategy based on a variety of multi-modal data inputs. The objectives aim at improving the quality of the interventions by making them easier, more accurate, and more intimately linked to pre-operative simulations where accurate objectives can be defined. In particular, one basic challenge is to guide a surgical tool according to a pre-planned strategy: to do so robots and 3D localizers (mechanical arms or optical sensors) perform real time tracking of surgical tools such as drills [6]. *AR* plays a central role in this domain because the key point of *CAS* systems is to "augment" the physical world of the surgeon (the operating theater, the patient, the tools etc.), by providing pre-operative information including the pre-planned strategy. Information is transmitted between the real world and the computer world using different devices: computer screen, mouse, pedal, tracking mechanism, robot, etc.

Since 1985, our laboratory is working on designing, developing and evaluating *CAS* systems. Through technological progress and a growing consciousness of the possibilities of real clinical improvements using a computer [18], *Augmented Reality* systems are now entering many surgical specialties. Such systems can take on the most varied forms [3]. Three classes of CAMI systems are identified in [19]:

- The passive systems allow the surgeon to compare the executed strategy with the planned one.
- The active systems perform subtasks of the strategy with the help of an autonomous robotic system.
- The semi-active or synergistic systems help the surgeon in performing the surgical strategy but the surgeon is in charge of its execution. The system and the surgeon are working in a synergistic way.

By comparison with the HCI domain, in which traditional design approaches aim at keeping the user in the loop and at focusing on the task supported by the system, *CAS* design methods are principally driven by technologies. Consequently, instead of assessing the quality of a system in terms of the user's perspective and of the software designer's perspective, *CAS* design methods are constrained and driven by technologies. External properties [1] [13], which establish "how usable a system is" and internal properties that describe the software quality, are replaced by clinical and technical considerations [17]. In this context, our research aims at providing a notation for designers to help them in reasoning about the merging of real and virtual entities.

More generally speaking, *Augmented Reality* systems design is often driven by the latest technology. We place a greater emphasis on interaction between the user and the system as well as the user and the real environment. To do so we propose a notation, namely ASUR: ASUR description of a system is composed of:

- the entities that are involved in the interactive system,
- the relation between these different entities.

In addition, characteristics are identified to describe the entities and the relations involved in the user's interaction with the system. Finally, ASUR provides a common description notation of *AR* systems that enables their comparison and their classification. The next paragraph presents this notation.

3. ASUR Interaction Description

In this paragraph, we first briefly present the principles of our ASUR notation. As mentioned above, ASUR is based on the principles of our previously presented OP-a-S notation [9]. ASUR extends OP-a-S by providing characteristics of components and relations involved in the dual interaction of the user with the virtual part of the system and with his/her real environment.

3.1. ASUR Principles Overview

The basic idea of ASUR is to describe an interactive system as a set of four kinds of entities, called components. In [12], we have already presented some characteristics of such entities, but the relations among them were not studied. When applying ASUR, a relation between two components describes an exchange of data. ASUR components and relations are described in the two following paragraphs.

3.1.1. ASUR Components

The first component is the **User** (component U) of the system. Second, the different parts used to save, retrieve and treat electronic data are referred to as the computer **System** (component S). This includes CPU, hardware and software aspects, storing devices, communication links. To take into consideration the use of real entities, we denote each real entity implicated in the interaction as a component R, **Real** objects. The 'Real object' component is refined into two kinds of components. The first component R_{tool} is a **Real** object used during the interaction as a **tool** that the user needs in order to perform her/his task. The second component R_{task} represents a real object that is the focus of the task, i.e. the **Real** object of the **task**. For example, in a writing task with an electronic board like the *MagicBoard* [7], the white board as well as the real pens constitute examples of components R_{tool} (real tool used to achieve the task), while the words and graphics drawn by the user constitute the component R_{task} (real object of the task). Finally, to bridge the gap between the virtual entities (component S) and the real world entities, composed of the user (component U) and of the real objects relevant to the task (components R_{task} and R_{tool}), we consider a last class of components called **Adapters** (component A). **Adapters** for **Input** (A_{in}) convey data from the real world to the virtual one (component S) while **Adapters** for **Output** (A_{out}) transfer data from the component S to the real world (components U, R_{tool} and R_{task}). Screens, projectors and head-mounted displays are examples of output adapters, while mice, keyboards and cameras may play the role of input adapters. The exchange of data between ASUR components is described in the next paragraph.

3.1.2. ASUR Relations

A relation is symbolized in an ASUR diagram with a unidirectional oriented arrow. It represents a set of data sent by a component to another one. For example, a relation $A_{out} \rightarrow U$, from a screen (component A_{out}) to a user (component U) symbolizes the fact that data are perceivable by the user on the screen. Another relation $U \rightarrow R_{tool}$, from a user (component U) to a pen of the *Magic Board* (component R_{tool}) represents the fact that the user handles the pen.

Having defined the ASUR components and relations, we now focus on the user and her/his interaction with the computer system as well as with the real environment.

3.2. Focus on the User's Interaction

Due to our definition of *AR* systems, the users' interaction has two facets: (1) interaction between the user and the computerized part, and (2) interaction between the user and the real environment. According to our ASUR notation, the first facet is represented by a relation from an output Adapter (A_{out}) to the User (U). Data from the computerized part may only be perceived through an output Adapter. Interaction between the user and the real environment (facet 2) is represented by ASUR relations between the component U (the User) and the components R_{task} (Real object of task) as well as the components R_{tool} (Real tool).

Getting a clear understanding of the interaction between a user and an *AR* system involves analysis of the following relations: $A_{out} \rightarrow U$, $R_{task} \leftrightarrow U$ and $R_{tool} \leftrightarrow U$. In the next two paragraphs, we characterize these components and relations.

3.2.1. Three Characteristics of an Adapter and a Real object

The first characteristic induced by the use of a real object or an adapter is the **human sense** involved in perceiving data from such components. The most common used ones are the haptic, visual and auditory senses. For example, in the *Magic Board*, the visual human sense characterizes the white board, the user looking at the drawings and written texts. The auditory sense may also be involved to perceive alarms indicating a problem for example with the vision-based capturing process.

The second characteristic of an adapter or a real object, is the **location** where the user has to focus with the required sense, in order to perceive the data provided by the adapter or to perceive the real entity. The coupling of the characteristic **human sense** with the **location** defines the **perceptual environment** of an adapter or a real object.

The last characteristic is the ability of the adapter or real object to simultaneously **share** the carried data among several users. For example, displaying data on a head-mounted display (HMD) restricts the perception to the user wearing the HMD. On the other hand, projecting data onto a white board enables *N* users to perceive the data simultaneously.

3.2.2. Two Characteristics of a Relation

An ASUR relation represents a flow of data. The **interaction language** used to express data carried by the relation is the first characteristic of a relation. The data transferred to the user may be expressed in an arbitrary manner [4]; e.g. the user needs to learn the form or syntax of the data. On the other hand, the language may be non-arbitrary; in this case, the data are expressed according to an already known convention. In the task of text selection on the *Magic Board*, the projector displays on the white board a square that follows the user's finger motions and delimitates the area of selection. The visualization of the selected area is thus non-arbitrary, since this is widely used in computer applications.

The second characteristic of a relation denotes the importance, for the user's task, of the data carried by the relation. Defined in [12] as the attention received, the **weight** of a relation is a continuous axis ranging from none to high. We keep three values: none, peripheral and high. During a writing task using the *Magic Board*, the white board receives much attention, while the camera and projector receive none. Weighting a relation enables the designer to identify the number of relevant data that the user must perceive during a given task.

Based on our characteristics, the analysis of the ASUR relations and components linked to the user (U) enables the designer to identify problems in the usability of a system. For example, too many relations heavy weighted and with adapters requiring different locations, may lead to difficulties for the user to perceive all the data useful for performing a task. In case of potential identified usability problems, it is important to notice that the designer could only change the characteristics of the adapters and their relations. Characteristics of the real entities (R_{task} and R_{tool}) and their relations to the user are prerequisites of the system. In the next paragraph, we show how predictive analysis of the consistency ergonomic property can be addressed using our characteristics of ASUR components and relations.

4. Predictive Analysis Based on ASUR

We base our predictive analysis process on the characteristics of ASUR components and relations as well as on ergonomic properties. As mentioned above, in *AR* systems, information or action are defined by the system to facilitate or to enrich the natural way the user would interact with the real environment. Outputs of the systems consequently constitute the focus of our analysis. Among the existing ergonomic properties, two of them are closely related to outputs of interactive systems, namely observability and honesty:

- **Observability** characterizes "the ability of the user to evaluate the internal state of the system from its perceivable representation" [8] [13];
- **Honesty** characterizes "the ability of the system to ensure that the user will correctly interpret perceived information and that the perceived information is correct with regards to the internal state of the system" [13]

Additionally Norman's Theory of Action [16] models part of the users' mental activities in terms of a perception step and of an interpretation step. The above two ergonomic properties are directly related to these two steps: Observability is related to the users' perception while honesty supports users' interpretation.

Observability and honesty are traditionally analyzed in the case of representation of one concept at a given time. Facing the formidable expansion of new technologies in the medical domain for example, the surgeon will be exposed to more and more sources of information: NMR data, Ultra Sound images, needle tracking data, etc. Observability and honesty of multiple concepts at a given time, and of one concept represented in multiple ways (representation multiplicity principle) must therefore be considered and involve consideration of another crucial ergonomic property: **consistency** across the variety of representations available at a given time. In the following table, we refine consistency in terms of **perceptual consistency** (observability level) and **cognitive consistency** (honesty level).

In terms of ASUR characteristics, **perceptual consistency** is ensured if every output adapter and real object that convey data to the user have:

- their corresponding **locations** compatible: their locations must spatially intersect,
- their associate **human senses** compatible: the user must be able to sense the different information without losing some of it.

In other words, perceptual consistency is established if every data conveyed by adapters and real objects, along heavily **weighted** relations, are simultaneously perceivable and do not imply that the user changes her/his focus of attention.

Addressing a problem of perceptual inconsistency requires selecting output adapters having their corresponding **locations** and **human senses** compatible with each other and with the ones associated with the real objects involved in the interaction.

	Perception (**Observability**)	Interpretation (**Honesty**)
1 concept, *n* representations	**Four**	**Types**
N concepts, *1* representation each	**Of**	**Consistency**

Table 1: Four types of consistency.

At the cognitive level, consistency extends the notion of honesty. In terms of ASUR, **cognitive consistency** is ensured if every relation from output Adapters (components A_{out}) or Real objects (components R_{tool} and R_{task}) to the User (component U) are based on the same **interaction language**. (see definition in 3.2.2). For example, data displayed by *KARMA* are based on a 3D graphical language (for example, arrows explaining which tray to open) matching the view of the real printer. In this example cognitive consistency is ensured because the view of the printer as well as the view of the 3D graphics match each other. If *KARMA* was displaying textual explanations, cognitive consistency would not be satisfied: indeed two languages are involved, a 3D view of the printer and a textual language.

Addressing a problem of cognitive inconsistency requires changing the **interaction languages** associated with the relations from the output adapters to the user.

During the design phase, if different solutions are designed, each solution envisioned is described with the same notation, allowing thus a precise comparison of the solutions in terms of consistency. The next paragraph illustrates our ASUR based analysis of consistency for one of our computer assisted surgery applications, *CASPER*. For more examples, we describe various *AR* systems using ASUR in [10].

5. ASUR Based Analysis: an Example

5.1. *CASPER* Application

5.1.1. Identity Card

CASPER (Computer ASsisted PERicardial puncture) is a system that we developed for computer assistance in pericardial punctures. The clinical problem is to insert a needle percutaneously in order to access the effusion with perfect control of the needle position and trajectory. The danger involves puncturing anatomical structures such as the liver or the heart itself. A detailed medical description of the system can be found in [5]. After having acquired Ultra-Sound images and planned a safe linear trajectory to reach the effusion, guidance is achieved through the use of an optical localizer that tracks the needle position. The left part of Figure 1 shows the application in use during the intervention.

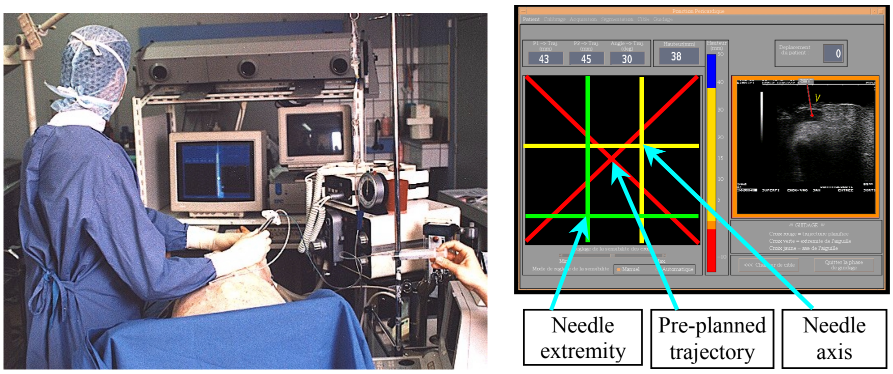

Fig. 1: Our *CASPER* application in use (left), *CASPER* guidance information monitored (right).

5.1.2. ASUR Description of *CASPER*

During the surgery, the surgeon (U) handles and observes a surgical needle (R_{tool}): **U↔R_{tool}**. The needle is tracked by an optical localizer (A_{in}): **R_{tool}→A_{in}**. Information captured by the localizer is transmitted to the system (S): **A_{in}→S**. The system then displays the current position and the pre-planned trajectory on a screen (A_{out}): **S→A_{out}**. The surgeon (P) can therefore perceive the information: **A_{out}→U**. Finally, the object of the task is the patient (R_{task}), who is linked to the needle, and perceived by the surgeon (P): **R_{task}→U**. Figure 2 presents the ASUR description of *CASPER*.

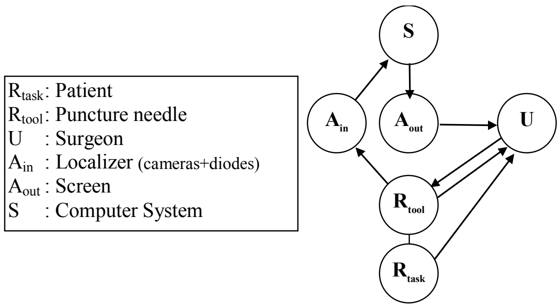

Fig. 2: ASUR description of *CASPER*.

5.2. Analysis of Cognitive Consistency in *CASPER*

An ASUR based analysis of *CASPER* has lead us to identify inconsistency between the perceived data from the needle and the ones displayed on screen. The arbitrary representation based on three crosses displayed on screen does not match the manipulation of the real needle. Since the puncture is a critical task that involves two relations that both require high attention ($R_{tool} \rightarrow U$ and $A_{out} \rightarrow U$), we deduce that cognitive consistency is not established.

To overcome this problem, we have worked on the way the guidance information is displayed. Instead of using the cross-based graphical representation, we have adopted a cone representation to visualize the trajectory. This design solution has brought up the problem of the point of view for displaying the 3D cone. Three points of view are possible: the needle or the trajectory point of view (the representation of the trajectory -respectively the needle - changes according to the position of the needle) or the user's point of view (the representation depends on the position and orientation of the user's gaze).

5.3. Analysis of Perceptual Consistency in *CASPER*

ASUR analysis of the output adapter (A_{out}) and real objects (R_{tool} and R_{task}) involved in the surgery lead us to identify inconsistency between the two locations associated with the screen (A_{out}) and the needle in the operating field (R_{tool} and R_{task}). While using *CASPER*, the surgeon must always shift between looking at the screen and looking at the patient and the needle.

To address this problem we have designed a new version of CASPER using another output adapter: a see-through head-mounted display (HMD). For cognitive consistency, we display the 3D representation of the trajectory from the user's point of view.

With the new version, the surgeon can see the operating field through the HMD as well as the guidance data in the same location. Additionally, we have used a clipping plane technique to perform a cut-away of a part of the trajectory representation, in order to match the depth of the real and virtual fields of view. Figure 3 presents the HMD we used and a view through the HMD under experimental conditions. We are currently performing acceptance tests with surgeons.

In order to assess the usability of the new version of *CASPER*, usability experiments are in progress in collaboration with colleagues of the Experimental Psychology laboratory. The goal of the experiment is to evaluate the usability of the new output adapter as well as the representation of the trajectory. We carried out the experiment with 12 participants. Each participant has to reproduce a predefined trajectory in the context of 8 different settings. The settings are defined by the device used to display the guidance information (screen or HMD), the representation of the guidance information (cone or crosses) and the point of view on this information (trajectory point of view or needle point of view). The first global outcome is an overall benefit resulting from the use of the HMD as compared to using the screen. This result is independent of the representation of the guidance information as well as the point of view on the information. Further analysis and results are awaited in the very near future, from data being currently analyzed.

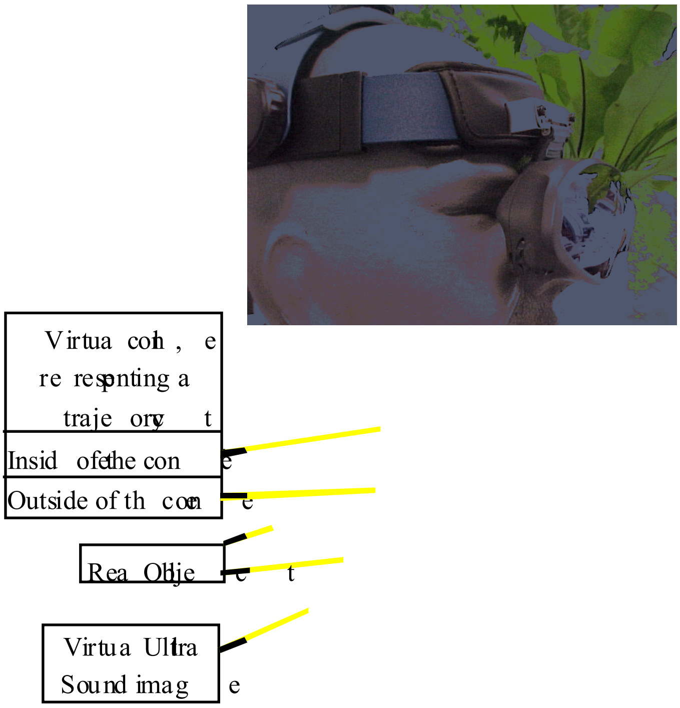

Fig. 3: The Sony head-mounted display used to address the perceptual inconsistency (top) and a view through the HMD merging real objects with a virtual conic trajectory (bottom).

6. Conclusion and Perspectives

In this paper, we have presented our ASUR notation and the characteristics of its components and relations as a tool to support predictive analysis of the interaction involved using *AR* systems. We showed that an ASUR description of a system could support the analysis of the consistency ergonomic property. The specificity of ASUR relies on the description of both real and virtual entities that are involved in performing a task in the real world. We illustrated our ASUR based analysis using *CASPER*, a *Computer Assisted Surgery* (*CAS*) system.

Although the inconsistency problems identified in the first version of *CASPER* can be detected without our ASUR analysis, we believe that ASUR provides a tool for systematically studying such usability problems and predicting several usability issues. In addition the simplicity of the ASUR description coupled with ergonomic properties makes it a useful tool for designers of *AR* systems and in particular designers of *CAS* systems who may not be familiar with ergonomic approaches.

In a future work, we plan to study other *AR* systems involving more complex information processes. We would also like to extend the number of ergonomic properties expressed in terms of ASUR in order to be able to cover a wider area of usability requirements.

7. Acknowledgements

This work is supported by the IMAG Institute of Grenoble and by the Grenoble University Hospital. We wish to thank the colleagues of the Experimental Psychology Laboratory of the University of Grenoble, who contribute to the new version of CASPER. Many thanks to G. Serghiou for reviewing the paper.

8. References

1. Abowd, D., Coutaz, J., Nigay, L., Structuring the Space of Interactive Properties, in Proceedings of Working Conference IFIP TC2/WG2.7 on Engineering for Human-Computer Interaction, (1992), p. 113-128.
2. Azuma, R., T., A survey of Augmented Reality, in Presence: Teleoperators and Virtual Environments 6, 4, (1997), p. 355-385.
3. Bainville, E., Chaffanjon, P., Cinquin, P., Computer generated visual assistance to a surgical operation: the retroperitoneoscopy, in Computers in Biology and Medicine, 25(2), (1995), p. 165-171.
4. Bernsen, N., O., Taxonomy of HCI Systems: State of the Art, ESPRIT BR GRACE, Deliverable 2.1, (1993).
5. Chavanon, O., et al., Computer ASsisted PERicardial punctures : animal feasability study, in Conference Proceedings of CVRMed/MRCAS'97, Grenoble, (1997), p. 285-291.
6. Cinquin, P., Bainville, E., Barbe, C., Bittar, E., Bouchard, V., Bricault, I., Champleboux, G., Chenin, M., Chevalier, L., Delnondedieu, Y., Desbat, L., Dessene, V., Hamadeh, A., Henry, D., Laieb, N., Lavallée, S., Lefebvre, J.M., Leitner, F., Menguy, Y., Padieu, F., Péria, O., Poyet, A., Promayon, M., Rouault, S., Sautot, P., Troccaz, J., Vassal, P., Computer Assisted Medical Interventions, in IEEE Engineering in Medicine and Biology 4, (1995), p. 254-263.
7. Crowley, J., Coutaz, J., Bérard, F., Things That See, in Communication of the ACM (CACM) Vol. 43 (3), (2000), p. 54-64.
8. Dix, A., Finlay, A., Abowd, G., Beale, R., Human-Computer Interaction, 2nd Edition, Prentice Hall, (1998).
9. Dubois, E., Nigay, L., Troccaz, J., Chavanon, O., Carrat, L., Classification Space for Augmented Surgery, an Augmented Reality Case Study, in Conference Proceedings of Interact'99, (1999), p. 353-359.
10. Dubois, E., Nigay, L., Troccaz, J., Un Regard Unificateur sur la Réalité Augmentée : Classification et Principes de Conception, in Revue d'Interaction Homme-Machine, Journal of Human-Machine Interaction, Europia Productions, (2001), 23 pages. To appear.
11. Feiner, S., MacIntyre, B., Seligmann, D., Knowledge-Based Augmented Reality, in Communication of the ACM n°7, (1993), p. 53-61.

12. Graham, T. C. N., Watts, L., Calvary, G., Coutaz, J., Dubois, E., Nigay, L., A Dimension Space for the Design of Interactive Systems within their Physical Environments, in Conference Proceedings of DIS'2000, (2000), p. 406-416.
13. Gram, C., Cockton, G., et al., Design Principles for Interactive Software, Chapman et Hall, (1996), 248 pages.
14. Ishii, H., Ullmer, B., Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, in Conference Proceedings of CHI'97, (1997), p. 234-241.
15. Médini, L. and Mackay, W.E., An augmented stripboard for air traffic control, Technical Report of Centre d'Études de la Navigation Aérienne, France, (1998).
16. Norman, D., A., Cognitive Engineering, User Centered Design, New Perspectives on Computer Interaction, Lawrence Erlbaum Associates, (1986), p. 31-61.
17. Notte, D., Nyssen, A., S., De Keyser, V., Evaluations des techniques de chirurgie minimale invasive par robot: premières constatations, in Conference Proceedings of ErgoIHM'2000, (2000), p. 234-243.
18. Taylor, R., Paul, H., Cutting, C., Mittlestadt, B., Hanson, W., Kazanzides, P., Musits, B., Kim, Y., Kalvin, A., Haddad, B., Khoramabadi, D., Larose, D., Augmentation of human precision in computer-integrated surgery, in Innovation and Technology in Biology and Medicine, Special Issue on Robotic Surgery, 13 (4), (1992), p. 450-468.
19. Troccaz, J., Peshkin, M., Davies, B., The use of localizers, robots and synergistic devices in CAS, in Conference Proceedings of MRCAS’97, (1997), p. 727-736.
20. Webster, A., Feiner, S., MacIntyre B., Massie, W., Krueger, T., Augmented Reality in Architectural Construction, Inspection, and Renovation, in Proceedings of Computing in Civil Engineering, ASCE, (1996), p. 913-919.