Assessing Continuity and Compatibility in Augmented Reality Systems

EMMANUEL DUBOIS [+], LAURENCE NIGAY [+1], JOCELYNE TROCCAZ [‡]

[+]*Dept of Computer Science, University of Glasgow, Glasgow G12 8QQ, Scotland*

(Emmanuel, Laurence)@dcs.gla.ac.uk
Phone: +44 (0)141 339 8855 (ext 0918)
Fax: +44 (0)141 330 4913

[‡] *TIMC-IMAG, Faculty of Medicine, (IAB), F-38706 La Tronche Cedex, France*

Jocelyne.Troccaz@imag.fr
Phone: +33 (0)4 76 54 95 08
Fax: +33 (0)4 76 54 95 55

Abstract. Integrating computer-based information into 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. Information is provided by the AR system to facilitate or to enrich the natural way the user interacts with the real environment. We focus on the outputs of such systems, so that additional information is smoothly integrated with the real environment of the user. We characterize the integration of the computer-provided entities with the real ones using two new properties: compatibility and continuity. After defining the two properties, we provide the elements and analytical method needed for assessing them. We also empirically study the two properties in order to highlight their impact on the interaction. The CASPER system, developed in our teams, is used to illustrate the discussion.

Keywords: Augmented Reality, Ergonomic Property, Continuity, Compatibility

Introduction

One of the recent design goals in Human Computer Interaction (HCI) has been to extend the sensory-motor capabilities of computer systems to combine real and computer-based information in order to assist the user in her/his environment. Such systems are called Augmented Reality (AR) systems. AR has been the subject of growing interest. In [1] several examples of AR systems are presented. Although AR systems are becoming more prevalent we still do not have a clear understanding of these systems. In [11], we emphasized the diversity of AR systems and presented one important classification characteristic:

---

[1] On sabbatical from the University of Grenoble, CLIPS Laboratory, BP 53, 38041 Grenoble Cedex 9, France.

- Systems that enhance interaction between the user and her/his real environment by providing additional capabilities and/or information. We call such systems, Augmented Reality ones.
- Systems that make use of real objects to enhance the interaction between a user and a computer. We call such systems, Augmented Virtuality ones.

In this paper, we focus on AR systems as defined above.

Within AR systems, our application domain is medical: Computer-Aided Medical Intervention (CAMI) systems. The main mission of CAMI systems is to help a surgeon in defining and executing an optimal surgical strategy based on a variety of multimodal data inputs. The objectives aim at improving the quality of the interventions by making them easier to perform, 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. Many CAMI systems have been designed and developed in many different surgical specialties. These systems support various types of interaction between their computerized parts and the surgeon, ranging from passive navigation or augmented reality to active robotic or tele-robotic assistance. Nevertheless, despite the existence of a number of experimental prototypes, relatively few systems have reached everyday clinical practice. One reason for this is that very little attention has been paid to software design and in particular to Human-Computer Interaction (HCI) design. Indeed the most attention has been paid to technical issues related to image processing, data fusion and surgeon assistance stemming from the clinical specifications of the problem. Very little effort has been applied to modeling and studying the interaction between the surgeon and the system. The design approach of current clinically oriented CAMI projects has been technology-driven [7]. But recently, the *Food and Drug Administration* recommended to take care about the end-user of CAMI system, the surgeons, when designing such systems. This recommendation has introduced the terms "*usability*" in this domain [2]. Usability is composed of four different aspects that have to be taken into account when designing CAMI systems:

- "Effectiveness": significant improvements for the patient should be provided when using the system, on a clinical point of view ;
- "Safety": the system must not present any risk during its use to the patient nor to the surgical team ;
- "Ease of use": the interactive system must be easy to use and to understand (intuitive) ;
- "Acceptance by the user": the user interface must be user-friendly.

Our approach is linked to the two last aspects identified by the FDA as part of the usability property. It is thus complementary to existing technology-driven approaches in that it focuses on interaction between the surgeon and the CAMI system. We focus our study of interaction between the surgeon and the CAMI system on the output user interface (from the system to the surgeon). Indeed, as previously explained, one of the main challenges of CAMI systems is to guide the surgeon according to a pre-planned strategy. Surgeon's actions are partially based on the outputs provided by the system. We therefore concentrate our analysis of interaction on the output user interface.

In this paper, we present our analysis of output user interfaces of AR systems. Our analysis is based on the two ergonomic properties: continuity and compatibility. First we define these two ergonomic properties that characterize outputs of AR systems. We then explain what are the characteristics of the interaction that must be considered in order to analyze those properties: i.e. the assessment elements. Having defined the properties and their assessment elements, we provide a method for assessing the properties that is based on our ASUR notation [12]. We illustrate our approach using our CAMI system: CASPER, a Computer ASsisted PERicardial puncture. In the context of CASPER, we illustrate our analytical method of continuity and compatibility and present our empirical studies of the two properties. We present the main features of CASPER in the next section.

An illustrative example: CASPER

CASPER (Computer ASsisted PERicardial puncture) is a system that we developed for computer assistance during pericardial punctures. The clinical problem is to insert a needle percutaneously in order to access an effusion in the pericardium, an envelope that surrounds the heart. The danger involves puncturing

3

anatomical structures such as the liver or the heart itself. Without this computer-assisted system, the puncture is performed using ultra-sound image control, leading to uncertainty about the position of the needle extremity and its orientation. CASPER provides real-time accurate information of the position and orientation of the needle, superimposed on a pre-planned trajectory. A detailed medical description of the system can be found in [6]. Basically, after having acquired a set of ultrasound images and planned a safe linear trajectory based on these medical images to reach the effusion, guidance is achieved thanks to the use of an optical localizer that tracks the needle position. The top part of Figure 1 shows the application in use during the intervention. The bottom part of Figure 1 shows the CASPER graphical user interface during the guidance phase. The interface consists of four areas hereafter referred to as visor, gauge, ultrasound image and numerical data. The ultrasound image and numerical data are not used by the surgeon during the surgery. The ultrasound image that is displayed, is one of the previously acquired images, whose plane of acquisition has been computed as being the mean of all acquisition planes. During the intervention the heart may be slightly deformed, for example by the pressure of the needle. Consequently, the surgeon's confidence in this pre-acquired image is not high. Numerical data represent distances in millimeters between the needle and the trajectory within the acquisition plane of the ultrasound image. Thus the surgeon's confidence in these data is again not high. In other words, the surgeon mainly relies on the gauge and the visor, while performing a surgery.

The gauge contains the different parts of the pre-planned trajectory: the skin and tissues, the effusion and the heart. In addition, a dynamic cursor is displayed on top of this static representation. The cursor denotes the current position of the extremity of the needle along the depth of the pre-planned trajectory. The cursor should thus never enter in the last part of the gauge (heart) and the effusion should be punctured only when the cursor is in the effusion part of the gauge.
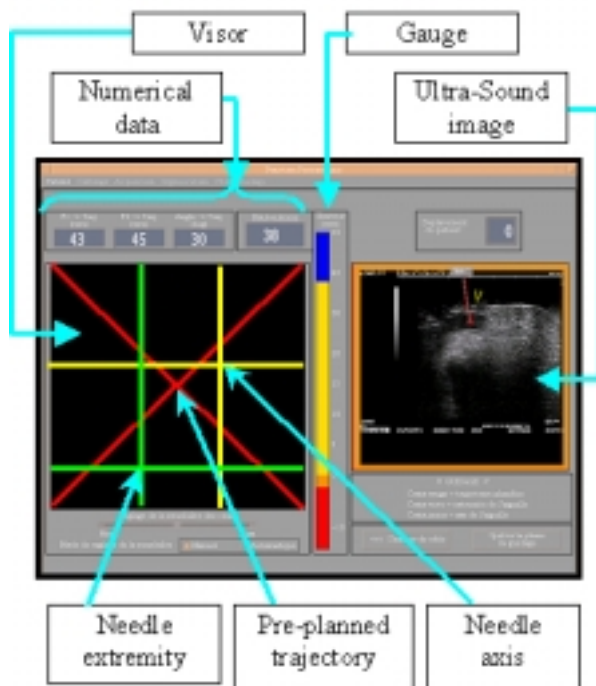
Fig. 1: CASPER application being used (top), CASPER screen, the guidance information (bottom).

Finally the visor is composed of three crosses:

- A red cross represents the pre-planned trajectory to follow and is a stationary cross ;
- A yellow cross encodes the position of a given point of the axis of the needle, according to the pre-planned trajectory to be reproduced ;
- A green cross corresponds to the position of the extremity of the needle, according to the pre-planned trajectory.

While the surgeon is inserting the needle through the body of the patient, the three crosses must be aligned. Indeed when the three crosses are superimposed the executed trajectory corresponds to the planned one. Moreover, to be sure that the surgical needle has not been distorted during its insertion, the surgeon regularly

checks the surgical needle. Consequently, during the introduction of the surgical needle, the surgeon has to look both at the guidance information displayed on the screen, and at the surgical needle in the operating field.

Continuity and compatibility

Definitions

As previously explained, one of the main challenges of CAMI systems is to guide the surgeon according to a pre-planned strategy. Surgeon's actions are partially based on the outputs provided by the system. We therefore base our analysis on the two ergonomic properties that characterize the output user interface:

- Observability characterizes "the ability of the system to ensure that the user can perceive the internal state of the system from the perceivable representation of that state". [9] [14];
- 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 regard to the internal state of the system" [14].

Transgression of the observability property makes realization of a task more difficult because the user does not have the information necessary to perform the task. Honesty is a more complex property to analyze because it depends on the expertise of the user. We illustrate this point with a simple example. Most CAMI systems include a registration stage that allows one to relate pre-operative data (from which the task is planned) to intra-operative data (on which the task is performed). For instance 3D points are collected intra-operatively by using a suitable pointer on the surface of an anatomical structure and are correlated with the segmented data. The way the surgeon collects data intra-operatively is very important: Indeed this registration stage is a key element of the procedure because the accuracy of the registration result has a direct impact on the accuracy with which the intervention will be performed. The accuracy of the registration result strongly depends on the collected data, the initial position guessed before registration, and the selected algorithm. It is also very important for the surgeon to evaluate the quality of the computed result before using it for the intervention. A key question related to observability and more closely to honesty is: How to

provide the surgeon with the result of registration so that s/he can evaluate it? Typical interfaces of experimental prototypes and even commercial systems propose numbers or percentages as basic information to be managed by the surgeon. The meaning of these numbers (interpretation process) is often obscure to the users since these numbers require them to have a good understanding of the algorithms: The honesty property is not verified. Other types of interaction may be selected: graphical interaction or test actions with the tool using the registration results. This last type of control is probably the best suited for the user in this clinical context.

With the large expansion of new technologies in the medical domain, the surgeon will be exposed to more and more sources of information: scanner data, echographic data, needle tracking data, etc. In this context, it is crucial to study what information must be presented to the surgeon and how to present the information according to the physical environment, i.e.: the surgery room, from which the surgeon also obtains information including the current distortion of the physical needle. These are two key properties. Due to the various sources of information from the computer as well as from the surgery room, we need to also consider the observability and honesty of multiple concepts at a given time. To do so, we distinguish the observability and honesty of multiple concepts (called compatibility) from the observability and honesty of multiple representations of a single concept. The latter case is called the continuity property. It is important to note that for both compatibility and continuity, some representations may be provided by the computer while some are physical such as a surgical tool. Additionally Norman's Theory of Action [17] models users' mental activities in terms of 7 steps, which include a perception step and an interpretation step. Observability and honesty ergonomic properties are thus directly related to these two steps: Observability is related to the users' perception (perceptual level) while honesty supports the users' interpretation (cognitive level). Consequently, as shown in Figure 2, compatibility and continuity can be applied at both the perceptual and cognitive levels:

- **Compatibility at the perceptual level** denotes how easy or difficult it is for the user to perceive all the concepts provided at a given time by the system.

- **Compatibility at the cognitive level** is assessed in terms of the cognitive processes involved in the interpretation of all the concepts perceived.
- **Continuity at the perceptual level** is verified if the user directly and smoothly perceives the different representations of a given concept.
- **Continuity at the cognitive level** is verified if the cognitive processes that are involved in the interpretation of the different perceived representations lead to a unique interpretation of the concept resulting from the combination of the interpreted perceived representations.

Perceptual Level	Observability	Perceptual Compatibility	Perceptual Continuity
Cognitive Level	Honesty	Cognitive Compatibility	Cognitive Continuity
	1 concept, 1 representation	n concepts, each 1 representation	1 concept, n representations

Fig. 2: Ergonomic properties: observability, honesty, compatibility and continuity.

Assessment elements

Having defined the two properties continuity and compatibility, we provide here elements that should be considered while assessing the properties. As for observability and honesty, we base our approach on the way information is conveyed to the user. Because we focus on Augmented Reality (AR) systems, we consider information provided by the system and information from the real world as being on equal footing.

Information from the system

First we consider the interaction modalities used by the system to convey the information. Our starting point is the definition of an interaction modality as the coupling of a physical device d with an interaction language L: <d, L> [15].

- A physical device is an artifact of the system that acquires (input device) or delivers (output device) information. Examples of output devices in CASPER include the screen.
- An interaction language defines a set of well-formed expressions (i.e., a conventional assembly of symbols) that convey meaning. The generation of a symbol, or a set of symbols, results from actions on physical

devices. In CASPER, examples of output interaction languages include the cross-based graphical representation of the guidance information.

In order to characterize the physical device, we define its perceptual environment as the coupling of the human sense used to perceive the information with the physical location where the user must focus her/his attention to perceive the information. For example, we will describe the screen in terms of the set <visual human sense, screen location>.

In order to characterize the language involved in the output interaction modality, we consider the cognitive resources a user will require in order to interpret the information. One characteristic would be the dimension or more exactly, the number of spatial dimensions of the representation that provide relevant information to the user. Defined in this way, the dimension of a language may be 1D, 2D, or 3D; Other characteristics of the language include those defined by Bernsen [4] in his theory of modalities: arbitrary, dynamic, analogue, linguistic.

Information from the real world

Second, we study how information from the real world is perceived and interpreted by the user. Indeed, in AR systems, such as CAMI systems, some relevant entities for performing a task belong to the physical environment. For example in CASPER, the surgical needle is a physical tool used to perform the puncture. By analogy with the information provided by the system, we identify two levels, the perceptual environment and the language. For example, checking if the needle is not distorted implies:

- the following perceptual environment : < visual human sense, operating field>
- a language being 3D and non-arbitrary [4] and involving particular cognitive resources a user will require in order to interpret the perceived information.

Continuity and compatibility assessment based on perceptual environment and language

When assessing perceptual continuity or compatibility, the perceptual environments involved in the interaction must be identified. Let us consider the assessment of perceptual compatibility. The various perceptual environments linked to the concepts that must be perceived in order to accomplish a given task must be considered. A perceptual incompatibility may be identified if the different environments make the perception difficult for the user: for example in a case requiring the user to look at different places. We rely here on results from theories of cognitive psychology such as ICS [3] and of perceptual psychology. For example, humans have the ability to selectively attend to one sound source in the presence of other sounds and background noise, and also listen to a background channel: the "cocktail party effect". Humans are constantly receiving a variety of information from the "periphery" without attending to it explicitly. Humans have highly sophisticated capacities for processing multiple information streams. As a consequence, two perceptual environments based on the sense of hearing does not systematically imply perceptual incompatibility.

While studying cognitive continuity or compatibility, the languages involved in the interaction must be identified. Again theories from cognitive psychology provide insights for studying the various languages. A general cognitive architecture such as ICS models the flow of information through different mental representations from sensation and perception, through comprehension, to action. This architecture also identifies cognitive aspects such as the influences of experience, memory requirements, and the potential for learning.

Because continuity refers to a single concept represented several times, there are more constraints among the participating languages than for the compatibility case. Indeed the user must be able to combine the various sensory representations in order to conclude that the representations correspond to the same concept. This fact is confirmed by cognitive psychology theories such as ICS. An ICS architecture constrains the way that different sensory representations (i.e., the user's "input" modalities) can be combined. Such a combination is not required

10

for compatibility because compatibility deals with different concepts perceived and interpreted at the same time.

Analytical method for assessing continuity and compatibility

In this section, we first present the ASUR notation [12] and then explain how ASUR can be used as a tool to evaluate the continuity and compatibility properties according to the assessment elements exposed in the previous section.

ASUR: a notation for describing the user's interaction

For a given task, ASUR describes an interactive system as a set of four kinds of entities, called components:

- Component A: Adapter (Input Adapter (Ain) or Output Adapter (Aout));
- Component S: computer System;
- Component U: User of the system;
- Component R: Real object involved in the task (tool (Rtool) or object of the task (Rtask)).

Subsequently a relation between two ASUR components describes an exchange between these two components. ASUR components and relations are described in the following two paragraphs.

ASUR components

The first component is the **User** (component U) of the system who interacts with the various components and benefits from their computational capabilities. Secondly, 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 involved 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 [8], where the digital and real ink are merged on a real whiteboard, the white board and 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 computer-provided 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 computer system (component S) while **Adapters** for **Output** (A_{out}) transfer data from 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.

ASUR relations

The interactive system is composed of these four components (A, S, U, R) which maintain relations with each other. We distinguish two types of relation in an ASUR description:

- A relation that stands for an exchange of data or an exchange of energy.
- A relation that denotes a physical link between objects.

On the one hand, exchange of data or energy is uni-directional and represented on ASUR diagrams with an arrow, from the source-component to the destination-component of the system. For example, a relation $A_{out} \rightarrow U$, from a screen (component A_{out}) to a user (component U) describes the fact that data, relevant to the task, are perceivable by the user on the screen. Another relation $U \rightarrow R_{tool}$, from a user (component U) to a pen of the MagicBoard (component R_{tool}) represents the fact that the user handles the pen.

On the other hand, a physical link between two objects is represented on ASUR diagrams with a single line between the two involved components. This type of relation occurs between two components of the real world (R_{tool} or R_{task}). This is

for example the case in the MagicBoard between the pen (component R_{tool}) and the white board (another component R_{tool}). Such physical links may also be observed between two adapters: the relation then represents a physical design constraint that combines two adapters into one. For example, in the Illuminating Light system [20], a camera (component A_{in}) and a projector (component A_{out}) are physically linked to form one single adapter, namely the IO-Bulb. Since this is a static permanent relation valid during the entire interaction, we have chosen to represent this last kind of physical link between two adapters ($A_{in} = A_{out}$) with a double line instead of a single line.

ASUR description of CASPER

To illustrate our notation, we describe here the ASUR modeling of our CASPER application. The task we consider for ASUR modeling is the realization of the puncture according to the pre-planned trajectory. Consequently the ultrasound images have already been acquired and treated and an ideal trajectory to puncture the effusion has been defined by the surgeon and saved on the computer system. In addition every pre-operative requirement, including the calibration of the needle, is done. The task, which we focus on, is thus limited to the introduction of the needle into the body of the patient. Figure 3 presents the ASUR description of CASPER. During the surgery, the surgeon is the user (component U). She/He handles and observes a surgical needle (component R_{tool}): $U \leftrightarrow R_{tool}$. The surgical needle is tracked by an optical localizer (component A_{in}): $R_{tool} \rightarrow A_{in}$. Information captured by the localizer is transmitted to the computer system (component S): $A_{in} \rightarrow S$. The computer system then displays the current position and the pre-planned trajectory on the screen (component A_{out}): $S \rightarrow A_{out}$. The surgeon (component U) is therefore able to perceive the information: $A_{out} \rightarrow U$. Finally, the person being operated, the patient, is not to be considered as a user (component U) but as the object of the task (component R_{task}). S/He is in contact with the needle ($R_{task} - R_{tool}$) and the surgeon sees the patient: $R_{task} \rightarrow U$.
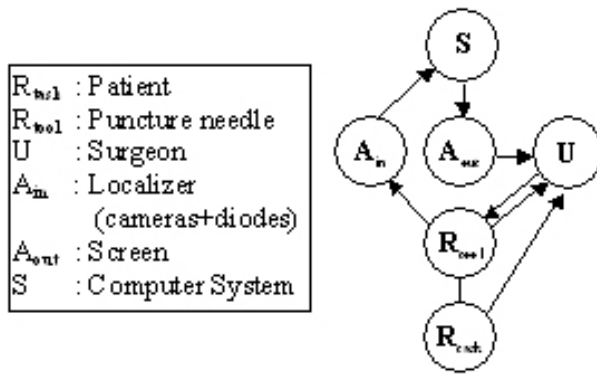
Fig. 3: ASUR description of CASPER.

ASUR-based analytical method for assessing continuity and compatibility

In an ASUR description, the user's interaction with the system is modeled by the relations connected to component U (User). The relations connected to component U represent the different facets of the output user interface. Based on these relations, we assess the two properties continuity and compatibility.

As explained in paragraph 3.2, discontinuity or incompatibility, at the perceptual level, is due to multiple perceptual environments <human sense, location>, that make the user's perception difficult. At the cognitive level, potential sources of discontinuity or incompatibility rely on the differences of the languages used to encode data relevant for the task. In an ASUR description:

- A perceptual environment characterizes a component which is linked to component U.
- A language describes a relation ending at component U.

In Figure 4, we summarize the characteristics of ASUR components or relations to be considered while assessing continuity and compatibility at the perceptual level as well as at the cognitive level.

Perceptual Level	Perceptual Environments of every component involved in the interaction	Perceptual Environments of components relevant to a given concept and involved in the interaction
Cognitive Level	Languages of every relation involved in the interaction	Languages of relations relevant to a given concept and involved in the interaction
	Compatibility	Continuity

Fig. 4: Characteristics of ASUR components and relations to be considered when assessing continuity or compatibility.

Based on Figure 4, we present a six step analytical method to assess continuity and compatibility at the perceptual and cognitive level:

- **First step**: Building an ASUR description of the interactive system in terms of components and relations for the task to be studied.
- **Second step**: Identifying the influential concepts involved in the realization of the considered task.
- **Third step**: Isolating the ASUR relations that are relevant to the task and that compose the output interface, i.e. relations ending at component U.
- **Fourth step**: Defining the perceptual environments of the ASUR components from which the relations identified during the third step start, i.e. the human sense they required so that the user may perceive the data and the location where these data are perceivable.
- **Fifth step**: Characterizing the languages of each ASUR relations identified during the third step.
- **Sixth step**: Assessing continuity and/or compatibility for the concepts identified in the second step, both at the perceptual and cognitive level by studying respectively the perceptual environments identified in the fourth step and the languages characterized in the fifth step (Figure 4).

This method has been applied and tested on several Augmented Reality systems [10]. In the following section, we illustrate the method using our CASPER system.

Continuity and compatibility: Analytical method applied to CASPER

1) Building an ASUR description.

The first step of our analytical method has already been exposed in paragraph 0. The ASUR description of CASPER is shown in Figure 3.

2) Identifying the influential concepts.

In CASPER several concepts are particularly important to the surgeon. The first one is the patient being operated. Another very important one is the surgical needle used to puncture the effusion, and more precisely its depth of penetration and its orientation in comparison with the pre-planned trajectory.

3) Isolating the ASUR relations that compose the output interface.

Based on Figure 3, these relations are:

- Rtask→U: The surgeon observes the patient.

- Rtool→U: The surgeon checks the surgical needle.

- Aout→U The surgeon reads the data displayed on the screen.

4) Defining the perceptual environments.

The location where one can observe the patient (component R_{task}) is limited to the operating field. Since the surgeon visually perceives information about the patient, the perceptual environment associated with this component is defined by the set **(vision, operating field)**. The needle (component R_{tool}) is perceived (visually and by touch) by the surgeon performing the operation. The perceptual environment associated with this second component is defined by the set **(vision / touch, operating field)**. Finally the screen (component A_{out}) requires the visual sense but, given the constraints of the context (sterile environment, limited space, etc.), the screen is positioned far from the operating field. Consequently the final perceptual environment required by this adapter is defined by the set **(vision, screen)**.

5) Characterizing the languages.

The first relation (R_{task}→U) and the second relation (R_{tool}→U) identified in the third step correspond to the perception of a real scene or object. The language is thus the natural 3D visual perception and is characterized by the set **(3D, non-arbitrary)**. The last relation (A_{out}→U) carries four forms of data (the visor, the gauge, the ultrasound image, the numerical data), described in section 2. As explained in section 2, the ultrasound image and the numerical data are not used for performing the task. We therefore study the two languages corresponding to the two forms: visor and gauge. The visor representing the trajectory and the needle is 2D and of arbitrary shape. Representation of the gauge is mono-dimensional. The gauge is expressed in a non-arbitrary manner since the movements of the cursor sketch the motion of the extremity of the needle along the pre-planned trajectory. Consequently, two languages are used by the relations A_{out}→U: **(2D, arbitrary,)** for the visor and **(1D, non-arbitrary)** for the gauge.

6) Assessing continuity and compatibility.

Continuity: Only one influential concept is present or represented several times: the needle (physical needle and two mobile crosses on screen). For the patient

though we only consider the patient and not her/his ultrasound representation during the surgery: i.e. the surgeon is looking at the patient during the surgery, whereas the ultrasound image (a representation of the patient) displayed on screen is not used by the surgeon during the surgery. We first study continuity at the perceptual level, considering the needle, an influential concept identified in the second step of the method. Based on the fourth step of the method, we identify a perceptual discontinuity due to the distinct locations required to visually perceive the real needle (location = operating field) and the one represented on screen (location = screen). Moreover the differences between the languages identified in the fifth step indicate cognitive discontinuity. Indeed the arbitrary representation based on two crosses displayed on screen does not match the manipulation of the real needle.

Compatibility: At the perceptual level, the two perceptual environments (vision, operating field) and (vision, screen) identified in the fourth step, make the user's perception of all the concepts relevant for the task (i.e., patient and pre-planned trajectory) difficult. Such differences in the location may be a source of perceptual incompatibility. In addition, the fifth step has concluded that different languages are in use while performing the puncture: the dimension of the data carried out is either 1D (the gauge), 2D (the crosses) or 3D (the patient). Furthermore, there is a mix of non-arbitrary representations (the reality and the gauge) and arbitrary ones (crosses). These differences constitute a source of cognitive incompatibility.

To sum up, discontinuity at the perceptual and cognitive level is identified in CASPER while focusing on the needle. In addition we identify various sources of incompatibility, at the perceptual level as well as at the cognitive level for the puncturing task. The next section presents design alternatives to address the problems of discontinuity linked to the needle. In addition, we report the results of a user experiment to evaluate the impact of perceptual and cognitive discontinuity on the behavior of the user.

Continuity and compatibility: Empirical studies of CASPER

Two experiments were performed in our laboratory on 12 subjects (7 men, 5 women), in collaboration with the Department of Experimental Psychology in

Grenoble (Laboratoire de Psychologie Expérimentale – LPE). These 12 subjects had no relationship with the medical or surgical domain but they had already been involved in previous user's experiments lead by the LPE and were "tagged" as dependant or independent. This "tag" is related to the way these persons establish spatial references in the real world. Dependant subjects rely preferably on visual and static information to get their bearing, while independent subjects get their bearing on the basis of gravito-inertial forces through their body in the real world. Outcomes of the experiments influenced by these psychological factors are not considered in this paper but are described in [10].

The task assigned to the subjects is similar to the CASPER puncturing task. Subjects were asked to reproduce a trajectory, stored in the computer, using a needle. A "Flock of Birds" [13] was used to localize the needle. Every 7 to 12 seconds, the computer emits a beep, signaling the subject to look at the real needle. The subject was also asked to confirm verbally that s/he has looked at the needle. The real conditions of the surgeon, who has to look alternatively at the surgical needle and the guidance information on screen, were thus reproduced. The main differences between the real intervention and the experimental conditions are the length and precision of the pre-planned trajectory to reproduce (about 5 cm long in an operating situation versus 50 cm in the experimental conditions) and the total absence of stress in the experimental conditions.

While performing the task, the difference (in x and y) between the position of the extremity of the needle and the pre-planned trajectory was recorded by the computer. Afterwards, these data were transformed into an energy spectrum, representing the energy spent by the user by oscillating around the pre-planned trajectory. A statistical ANOVA analysis of the collected data was performed with a threshold p = 0,05.

The subjects performed the task in different situations. In each situation, the interaction was different in order to study the discontinuity problem identified in CASPER. We present these different situations and corresponding results in the next paragraphs.

Addressing the perceptual discontinuity

Perceptual discontinuity is due to the perception of the needle in two different perceptual environments: (vision, operating field) and (vision, screen). To address this problem, the system must be able to enable the user to perceive the real needle and the computer-provided needle (two crosses) according to the pre-planned trajectory at the same location. Selecting the screen to both display the pre-planned trajectory and a video of the real needle is one solution to address the discontinuity problem. Nevertheless, these display options imply that the surgeon will operate without looking at the patient and the operating field: Such a solution must be carefully designed for safety and ethical reasons. Alternatively, guidance information could be displayed on top of the patient. "Image Overlay"-like solutions [5, 18] have been developed but remain cumbersome in an operating theatre. The guidance information could also be provided through tactile feedback [19]. But most of the tactile-based system also relies on complementary visual feedback. Finally, the solution that we have implemented is based on a see-through Head-Mounted Display (HMD). The surgeon perceives the guidance information in the HMD and perceives the real world i.e. the operating field, through the HMD.

Thus, in the first experiment, all of the subjects have realized the task with both display devices (screen and HMD) in a random order. The trajectory and the needle were represented by the CASPER visor and gauge (Section 2, Figure 1). The statistical analysis of the data collected during the 24 attempts has lead to the identification of a significant effect linked to the device, in favor of the HMD, with $p < 0.0012$ and a Fischer ratio (F) of more than 20. The F (Fisher) ratio is the ratio of the variance between groups ("treatment effect") to the variance within sample groups ("inherent variance") and is the basis for ANOVA. Figure 5 shows the mean obtained in each situation and highlights the fact that the HMD implies a smaller energy consumption by the subject than the screen. This result underscores the impact of perceptual continuity on the interaction.
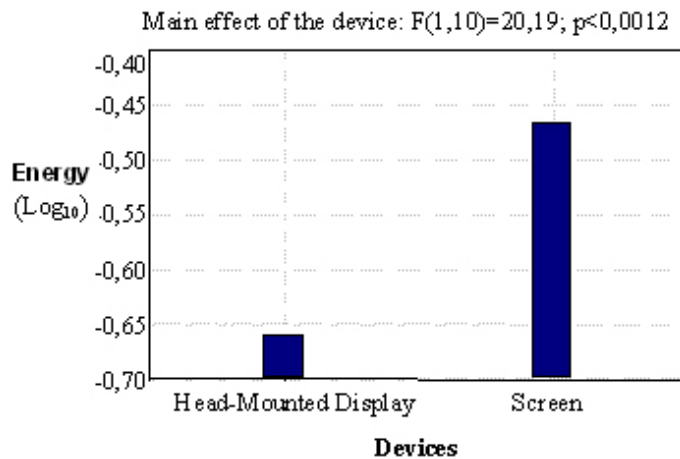
Fig. 5: Energy spent by the user as function of the devices used to display the guidance information.

Addressing the cognitive discontinuity

As explained in section 5, cognitive discontinuity can be studied by considering the multiple languages involved in expressing data related to the same concept: the needle. On the one hand, the perception of the real needle is based on 3D vision. On the other hand, the visor and the gauge representing the needle according to the pre-planned trajectory on screen correspond to two languages, a 2D and arbitrary language (crosses) and a 1D and non-arbitrary language (a rule). In order to address the cognitive discontinuity, without changing the real needle, we modified the representation of the guidance information, namely the visor and the gauge.

The solution, that we developed and proposed to the subjects during the second experiment, consists of a 3D representation containing the visor and the gauge. We represent the visor as a cone. Its summit represents the target, an ideal point in the effusion to bring the extremity of the needle in order to puncture the effusion. The center of the cone represents the pre-planned trajectory. The height of the cone stands for the length of the displacement of the needle from the patient's skin to the inside of the effusion, while the width of its base represents the initial accepted tolerance. The width of the cone decreases as the needle moves along the cone: this symbolizes the reduction of the accepted tolerance as the needle is inserted deeper and deeper into the patient. Moreover, in order to present the guidance data in a similar manner as in the previous version of CASPER, it is necessary to also add a 3D representation of the needle inside the cone. When

20

using this setting, the user perceives a static cone and the 3D model of the needle that moves according to the displacements of the real needle. The 3D representation of the needle should never go out of the cone to guarantee that the trajectory has been correctly reproduced. Finally, to show the progression of the needle along the trajectory, the point of view on the cone is moved along the height of the cone. The information previously presented by the gauge is now presented as a texture on the inner surface of the cone, enabling the user to perceive the depth of penetration and the orientation in a single 3D representation. Figure 6 shows a view of the guidance information displayed in this setting.
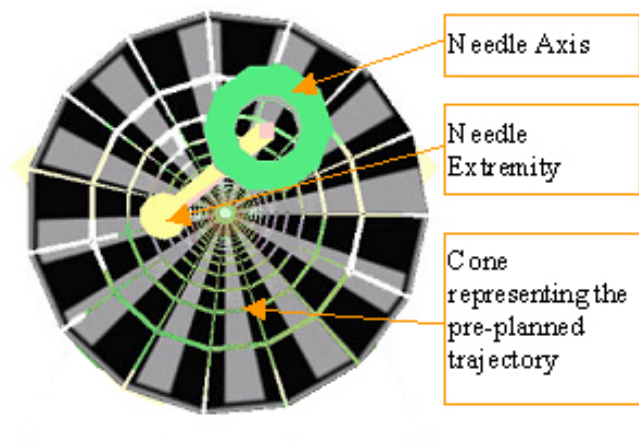

Fig. 6: 3D representation of the needle and 3D cone representing the pre-planned trajectory.

The second experiment aimed at comparing the use of the visor and gauge with the use of the 3D representation of the trajectory and the needle of Figure 6. Once again, subjects have to reproduce the pre-planned trajectory with these two settings in a random order. In both cases, the device used to display the information was the screen. The statistical analysis of the data collected during those 24 new attempts has lead to the identification of a significant effect linked to the representation, in favor of the cone, with $p < 0.0001$ and a Fischer ratio (F) of more than 46. Figure 7 shows the mean obtained in the two cases and highlights the fact that the cone representation implies a smaller energy consumption by the subject than the visor and gauge representation. This result underscores the importance of cognitive continuity on the interaction.
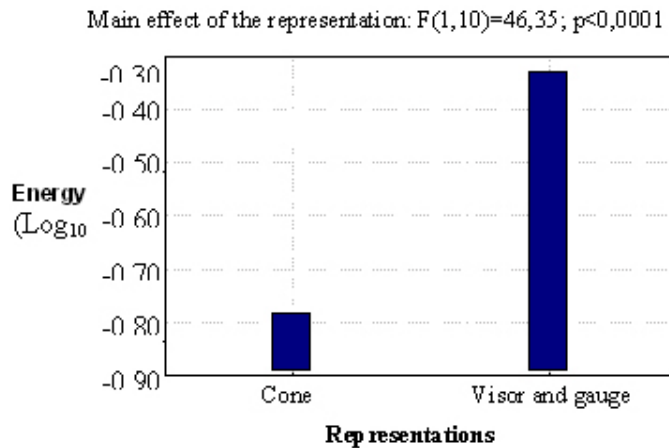
Fig. 7: Energy spent by the user as function of the representations of the guidance information.

Combining perceptual and cognitive continuity

The above very encouraging results lead us to consider the merging of the two design alternatives presented and tested in the two previous paragraphs. Using simultaneously an HMD with a representation based on a cone would verify the perceptual and cognitive continuity. The energy consumption should thus be less than in the above two experiments. However with this solution, the user has to perceive the real needle (by definition from the user's point of view), while the guidance information is represented from the point of view of the pre-planned trajectory. It might thus be difficult for the user to perceive simultaneously superimposed information expressed from two different points of view. Of course, the best solution would be to align the guidance information displayed in the HMD with the real world. However perceiving the trajectory directly matched on top of the patient may raise the problem of accuracy due to the size. It is difficult to perceive in the HMD a cone of only one millimeter width in which the user must insert a real needle. Further experiments will be conducted to evaluate this feature. Nevertheless, this discussion underlines the fact that the point of view chosen for representing data is a characteristic of the language that must also be considered.

Conclusion and Perspectives

In this paper we focused on two ergonomic properties of Augmented Reality systems: continuity and compatibility both at the perceptual and cognitive level. These properties focus on how information from the computer as well as from the

real world is smoothly perceived and interpreted by the user. We provide elements and an analytical method to assess these two properties while designing a system. The usability and usefulness of the method have been demonstrated in the context of our CASPER system: In particular the application of the method led us to design alternative solutions that have been developed in the new version of CASPER. Complementary to the analytical method, we also conducted empirical studies that show the impacts of the two properties on the interaction.

Further work needs to be carried out to better understand the characteristics of the language that are sources of cognitive discontinuity and/or incompatibility. For example, in the previous section, we sketched the importance of the point of view imposed by the language. Further collaborations with psychologists will be necessary in order:

- to identify new characteristics of the language,
- to experimentally evaluate the impacts of those characteristics on the interaction,
- to theoretically explain the impacts of those characteristics on the interaction.

Finally, we would like to point out the generality of the two properties, continuity and compatibility, that can be applied to any kind of interactive systems (not only Augmented Reality systems). Indeed, the two properties can for example be studied for two different representations on screen: In [16] we define a design rule called "spatial continuity" between two visual representations on screen.

Acknowledgements

References

1. Azuma, R., A survey of Augmented Reality. In Presence: Tele-operators and Virtual Environments, 6, 4, (1997), p. 355-385.

2. Babbitt, B., "*Medical Device Usability Engineering*", Workshop of the conference MMVR'2001, Newport Beach - CA, USA, (2001).

3. Barnard, P., May, J., Cognitive modeling for user requirements. Computers. In Communication and Usability: Design issues, research and methods for integrated service, Byerley, P., Barnard, P., May, J. Eds., Amsterdam: Elsevier, North Holland Series in Telecommunication, (1993), p.101-145.

4. Bersen, N., O., Taxonomy of HCI Systems: State of the Art. ESPRIT BR GRACE, deliverable 2.1, (1993).

5. Blackwell, M., Nikou, C., DiGioia, A., Kanade, T., An Image Overlay System for Medical Data Visualization. In Conference Proceedings of MICCAI'98, (1998), p. 232-240.

6. Chavanon, O., et al., Computer ASsisted PERicardial punctures: animal feasability study. In Conference Proceedings of CVRMed/MRCAS'97, (1997), p. 285-291.

7. Cinquin, P., et al., Computer Assisted Medical Interventions. In IEEE Engineering in Medicine and Biology, 4, (1995), p. 254-263.

8. Crowley, J., Coutaz, J., Bérard, F., Things That See. In Communication of the ACM, 43, 3, (2000), p. 54-64.

9. Dix, A., Finlay, A., Abowd, G., Beale, R., Human-Computer Interaction, 2nd Edition, Prentice Hall, (1998).

10. Dubois, E., Chirurgie Augmentée : un Cas de Réalité Augmentée ; Conception et Réalisation Centrées sur l'Utilisateur. PhD dissertation, University of Grenoble, France, (July 2001), 275 pages.

11. 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, Sasse A. & Johnson C. Eds, IFIP IOS Press Publ., (1999), p. 353-359.

12. Dubois, E., Nigay, L., Troccaz, J., Consistency in Augmented Reality Systems. Engineering for Human-Computer Interaction, Little R. & Nigay L. Eds, Springer-Verlag Publ., LNCS 2254, p. 117-130.

13. Flock of Birds, Technical description. http://www.ascension-tech.com/products/flockofbirds/

14. Gram, C., Cockton, G., et al., Design Principles for Interactive Software, Chapman et Hall, (1996), 248 pages.

15. Nigay, L., Coutaz, J., A Generic Platform for Addressing the Multimodal Challenge. In Conference Proceedings of CHI'95, Katz I., Mack R., Marks L. Eds., ACM Press: New York Publ., (1995), p. 98-105.

16. Nigay, L., Vernier, F., Design method of interaction techniques for large information spaces. In Conference Proceedings of AVI'98, Catarci, T., Costabile, M., Santucci, G., Tarantino, L. Eds, ACM Press Publ., 1998, p. 37-46.

17. Norman, D., Cognitive Engineering, User Centered Design, New Perspectives on Computer Interaction, Lawrence Erlbaum Associates, (1986), p. 31-61.

18. Peuchot, B., Tanguy, A., Eude, M., Virtual Reality as an Operative Tool During Scoliosis Surgery. In Conference Proceedings of CVRMed'95, (1995), p. 549-554.

19. Schneider, O., Troccaz, J., Chavanon, O., Blin, D., Synergistic robotic assistance in cardiac procedure. In Conference Proceedings of CAR'99, (1999), p. 803-807.

20. Underkoffler, J., Ishii, H., Illuminating Light: An Optical Design Tool with a Luminous-Tangible Interface. . In Conference Proceedings of CHI'98, ACM Press: New York Publ., (1998), p. 542-549.