A Mixed Reality System for Design Engineering: Thinking, Issues and Solutions

Nelly de Bonnefoy

EADS Corporate Research Centre
Toulouse Computer Science Research Institute
(I.R.I.T.)
31700 Blagnac, France
33(0)5 61 58 48 51
Nelly.de-bonnefoy@eads.net

Jean-Pierre JESSEL

Toulouse Computer Science Research Institute
(I.R.I.T.)
Paul Sabatier University
31400 Toulouse, France
33(0)5.61.55.63.11
jessel@irit.fr

ABSTRACT

Despite the extended use of Digital Mock-Up (DMU) applications during the development of complex aeronautical products, some major changes can be requested in order to correct deficiencies or to improve further products performances, especially when the first one comes out production. These changes are discussed upon the current product configuration and imply trades-off between different competencies involved in the product's lifecycle. Based on a specific use case, we first highlight the underlying issues associated to the development and introduction of Mixed Reality (MR) systems. We tackle these issues from cognitive and ergonomic points of view as well as technical. Then, we propose several MR solutions, which will be studied further during the two remaining vears of our research project. Theses solutions will then be evaluated against the specific needs and requirements of the aeronautic industry.

Keywords

Mixed Reality, Interactions, Design Review.

INTRODUCTION

When the first product comes out production during the development of complex aeronautical products, to correct deficiencies or to improve further products performances, some major changes can be requested. They are discussed upon the current product configuration. These changes also imply trades-off between different competencies involved in the product's lifecycle. It is typically a collaborative work situation where a team of designers gather around a table to perform a product design review. Information sharing and negotiation movements during those review sessions are strongly influenced by the respective speciality, knowledge and experience of participants. Information

exchanges around the physical object aims at taking some decisions concerning the future product configurations such as systems segregation, ergonomics, and physical arrangement of components. Consequently, during those reviews, debates between the different designers can lead to a request for an engineering change of the product.

Usually, information exchanges are made verbally helped with hand made sketches. In those cases, it might be difficult for a designer to make his engineering change proposal understood by others team members. This situation highlights a requirement implied by this kind of meeting: the need to visualise the proposed modification, in order to assess the potential changes impacts.

Mixed Reality (MR) systems can efficiency support interparticipants information sharing around a physical object. As matter of fact, MR systems can bring some more engineering information about the physical object through "optical see through"-based visualization applications. Part changes are made visible by adding complementary information to the current part («Augmented reality» concepts)(Cf. figure1). This "augmentation of reality" can be realised with the image incrustation in the users field of view. Incrustation are performed through, optical seethrough head mounted displays. Those display devices allow users to keep their environmental perception while having intuitively access to more contextual information. Therefore, such systems leverage the typical limitations of paper-based systems, and add value to exchanges and tradeoff through of their intuitive and fast access mode to pertinent information.

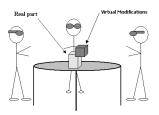


Figure 1 - The Context

SYSTEM DESCRIPTION

The investigated system allows users to virtually simulate the integration of an engineering change by modifying designers' perception of the real object.

During the meeting, all participants share the visualisation of the actions performed on the augmented model, but the access to the model itself must remain. So that conflicts can be avoided. Practically, a designer must have the knack if he desires to create virtual modifications. One design sequence could be performed like this: he selects a virtual geometric form in one menu, it displays in front of him, and he can modify its attributes by using the sub-menu associated to the virtual form. He is free to place the virtual modification of the real object when and where he wants.

The introduction of such systems within the review environment must preserve natural communications and social cues between participants. Indeed, the review process and user behaviour must not be perturbed by a complex MR system and its multiple accessories. The use of such system has to be as natural and intuitive as possible, this approach is similar to the "Natural User Interface" (NUI) developed by [1].

This system comes also under the definition of Collaborative Mixed Reality systems. Some references already exist in the literature:

- "Build-it": a collaborative mixed reality system which tend to be a NUI [2],
- "Arthur": the development of an interactive task oriented collaboration environment based on augmented reality [3],
- "MARE": a multi-users augmented reality environment on table set-up [4].

INFORMATION SHARING

Information diversity

Different types of information exchanges take place during review meetings, the next section focuses on system/user information transactions.

We considered two major typologies of information to be manipulated by the user: textual information, and graphical information. Both type can be broken down as follows:

Text information of consists of:

- interface information (menus, etc.),
- system messages,
- annotations the user wants to add.

Graphical Information consists of:

- arrows, chips, or others in order to show something to advantage,
- classic geometric shape the user will manipulate in order to obtain the final modification,

• the final change.

Depending on the nature of the information to be communicated to the user, relevant output interaction modalities have been identified. The next section present the various output modalities used and for each, the different types of information associated to them, and the best user-centred utilization.

Output modalities

Three different output modalities have been considered in order to make more perceptive actions:

- the visualisation through a Head Mounted Display (HMD),
- the text-to-speech system,
- and the haptic force feedback.

Display of the HMD

The optical see-through HMD is a display device allowing users to keep their environment perception while visualising more information in their field of view, in context.

Hardware

The figure 2 describes the basic principle of the optical seethrough Head Mounted Display:

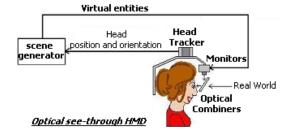


Figure 2 - Optical see-through HMD

There are two main devices interesting, the Sony Glasstron (no more on the market) and the Nomad from Microvision.

Equipped with this device, each designer visualizes all information in his field of vision, such as his environmental information and system information. The system must perform display in the most user-friendly and ergonomic way so the user cannot be lost within the information display density.

For that, we discuss the general exploitation of the field of view and we detail requirements display for each kind of information.

Usually, people fields of view consist of one main area that is a zone in the sight direction where people see all things clearly, and around it, the field of view area where people have to move eyes to see distinctly information (Cf. Figure 3):

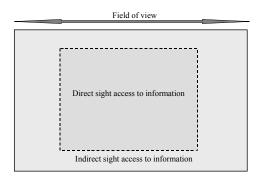


Figure 3 - Field of view organisation

One can be noticed that, most of the time, when people have to look at something on the right (or left), even if it is close to them, the head moves more than eyes. People tend to put the subject of their attention in the centre of their field of view.

Each type of information does not require the same display modes. Displays do not have the same duration, neither the same location imperatives (Cf. Table 1).

	LOCATION		
	Specific area, Link with part	Specific area, no link with part	No specific area, no link with part
Interface (1)		X	
Syst messag. (2)		X	
Annotation (3)	X		X
Arrows (4)	X		
Geom. Shape (5)	X		X
Final change (6)	X		

Table 1 - Information type location

These information displays are developed hereafter.

Interface information

In order to get designers used to the location of interface information (1), specific zones are associated to this type of information (Cf. Figure 4). It is located as follows:

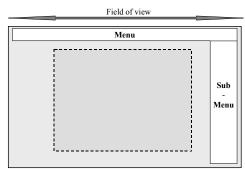


Figure 4 – Interface, menu and information location

Interface information are in this area because users do not have to focus their attention upon it. Users do not need to keep this type of information at time in the centre of their field of view. This area is visible only if the user has the knack.

System message

System messages (2) are important for users to navigate within the application. Users must perceive this kind of information as quick as possible. That is why they will be displayed in the centre of their field of view. A special colour has to be selected, in order to increase messages ergonomic and cognitive values.

Annotation and geometric shape

At the beginning of the process, annotation (3) and geometric shapes (5) do not need, necessarily, to be connected with the real object. The system will display this kind of information in front of the user field of view, in the "direct sight access to information" area. They could be modified, moved, deleted in the users own way.

Final change and Arrows, chips, etc.

The information category that concerned final change (6) and arrows, chips, etc (4), is used to emphases particular items. For example, it will be used to precise a place through an arrow or to show the real part modification through a virtual change.

Their location in the users' environment is important. They depend on the real part location in the meeting room and not on one specific user field of view area. The modification part, which is geometric shape, has an exact location on the real object; however, arrows and others only require an accurate location associated to the real object.

General

Different colours will be used: each designer will have his own colour, in order to differentiate users' actions.

Different colours are associated to system information, one for menus and one for messages. The selected menu section will have different background than others, and if the user goes down in sub-menus the navigation path will be displayed as a key words list.

Moreover, users have the opportunity to remove display; all visual information will be hidden if he desires so.

Text to speech

This output modality is the sense of hearing one.

The hardware is a headphone.

The most important part, with this modality is to have a good thought concerning its use. If text-to-speech is used permanently, for all information in text form, it becomes unbearable for users very fast. And in that case, users put the sound off.

This modality has to be used only if it gives an added value to the system situation. That is why it is generally fitted for classic system messages: system announcement or error messages. Users' attention has to be focused on this type of information.

Haptic force feedback

This output modality is used to modify, and move the geometric shape that will become the final change and also to touch the final change. As matter of fact, the haptic force feedback is very important to manipulate virtual objects. To make the system more realistic, users must perceive tactile information. But today, without special devices, it is almost impossible for users to catch this information from virtual objects.

Hardware

There are two main different accessories: the glove and the Phantom, which is a computer device most closely related to the mouse. Their function is to interact with objects in a three dimensional environment. During the last few years, many research have been carried out in this domain, and accessories progresses in ergonomics are significant.

There is, for example, the Cyberglove (Cf. Figure5) with a vibrotactile feedback [5]. There are small vibrotactile stimulators on each finger and the palm of the cyberglove. They can produce complex tactile feedback patterns. Even if this glove is an accessory, the ergonomic material aspect is studied: users cannot feel that they have a robot hand, they just wear an ordinary glove.



Figure 5 – Cyberglove [6]

Because the system has to be as intuitive as possible and because there is no method without accessories, ergonomics issues play a major role. The devices that will be selected must allow users to keep their usual meeting behaviours. For example, they should catch objects if they want, such as a pen, a glass, etc.

In an information exchange there are two communication ways, in our case: "system (S) to user (U)" and "user to system". The last section explains how the user perceives information from the system (S \rightarrow U). The next section deals with the User to System communication (U \rightarrow S), which are kinds of interaction the user can perform.

INTERACTION - INPUT MODALITIES

Human-system interactions have the objective of developing models, concepts, tools and methods, in order to realise systems that answer users' needs and aptitude. Reproducing usual human-human communication modes, the modalities used in this system are: voice and gestures.

This choice has been made because interactions with the system must remain as intuitive as possible.

As we already mentioned there are different types of information to be manipulated. The user has to interact with all of them.

Speech recognition

In this system, the integration of speech recognition is done in two stages.

First, speech recognition is used to browse menus and submenus. In this case, the system must be able to recognize words rather than sentences. Commercial off-the-shelf applications are enough efficient to perform these functions. But a great deal of attention must be paid to the design of menus and to the selection of a clear and concise vocabulary.

In the second integration stage, the speech recognition will be used to navigate in menus, to modify virtual changes and integrate them on the real product. This will imply that the system will perform sentence recognition making the system more friendly and intuitive.

Gestures recognition

Like speeches, gestures are a spontaneous mean for people to communicate with other actors. The use of gestures in multimodal applications facilitates users interactions, in particular in noisy environments. Moreover, users tend to execute gestures for manipulation operations rather than state them or access to them by using classical interfaces like window, icon, etc.

The system use gestures recognition to interact with interface information and to modify the shape of the changed geometric form. This feature, in some cases, will be used simultaneously with speech recognition features. The goal is to identify and track the gestures of the user that has the knack. As people know, there are different ways to make gestures recognition.

Method with Digital Gloves

Flexion angles measurements, which are obtained with an optic fibre positioned on each finger, give fingers configuration and position. These angles are determined by the luminous signal intensity sent in the fibre and with its intensity in tip of finger. A tracker is located on the hand in order to process the hand position and orientation. This method gives accurate results, but it constraints users to wear a glove generally linked to a system (depending on technologies employed). Users do not have hands free.

Visual Methods [6] [7]

These methods are based on computer vision and on image processing techniques. Hands movements are recorded with one or more video cameras. Then different techniques can be used to process images depending on the gestures recognition method used. It is more difficult to use this kind

of method but users get rid of physical accessories. Most of the processing techniques consist of four operations, which are acquisition, segmentation, characteristics extraction and classification. They can be realised in different ways: based on markers, on three-dimension model, on visual appearance. The main advantage of these methods is that users do not have to wear physical artefacts.

This input modality will be use to navigate in interface menu and to modify and move virtual geometric shape.

Now that all modalities have been defined, the tracking system, which has a primordial role in Mixed Reality System, must be tackled.

TRACKING AND REGISTRATION

To perform an incrustation of a virtual object in the user field of view, scene components need to be located accurately. Indeed, to make a good registration, MR systems need trackers with approximately one millimetre accuracy in position and a low fraction of degree in orientation. Most of commercially available trackers answer one of the two conditions but not both.

Tracked elements

In order to offer an intuitive and a free visualisation of part modification, the system will track continuously and with accuracy different environment elements:

- the physical part,
- each designer,
- users points of view,
- users hands.

The first three elements are tracked to make an efficient registration of the virtual modification on the real object in each user field of vision. The last one is tracked so users can realize virtual changes with gesture recognition, move the virtual object, and perform haptic force feedback.

The tracking system is designed to create relationships between each tracked element configuration (Cf. figure 6).

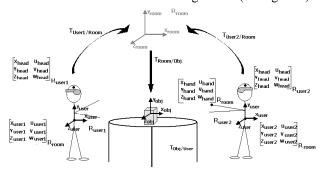


Figure 6 - Referential example

As people know, there are different sorts of trackers:

- Electromagnetic trackers (alternating current, direct current, compass)
- Acoustic trackers (distance measurements determined by ultrasonic time of flight),
- Optical trackers (with punctual receptors (phototransistors), or video based tracking),
- Mechanical trackers (inclinometer, gyroscope, accelerometer...),
- GPS trackers.

The good configuration of trackers has to be found. But some considerations have to be taken into account:

In order to make our system as natural as possible, the use of peripheral is limited and bulky peripheral are proscribed. The less the system uses devices, the better it is.

For instance, as one device used for the haptic force feedback, could be also used for gestures recognition as well.

Moreover, this use case takes place in a specific meeting room that is a well-defined place and where the luminosity is constant.

Considering the huge advances made in video technologies during the last few years, it is now possible to find little cameras with very good resolution. This greatly improves the quality of image processing for markers recognition. Moreover these video cameras are now equipped with USB communication port, providing a good data quality and speed transfer. Practically all web cams have these characteristics today, so a good camera could be found for a reasonable price. To finish with, little cameras start to be equipped of IEEE communication port, which offers the best transfer speed and quality.

For all these, the video based tracking has been selected for our system. Possible video based tracking methods are presented in the next section.

Video based tracking methods [8]

As people know there are two main video-tracking configurations:

• Inside out (Cf. Figure 7)

One or many video cameras are on the moving target. They watch markers fixed in the environment which are references.

Outside in (Cf. Figure 8)

One or many video cameras are fixed in the environment, they are references; they watch the target movements, on which markers have been affixed.

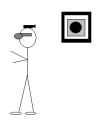




Figure 7 - Inside-out

Figure 8 - Outside-in

Markers should be classic draws or LEDs. Once the configuration is chosen, there are different manners to calculate the target location. The first uses the two or more cameras in order to calculate the target marker positions, by using the triangulation for example. The target orientation should be calculated by using several markers on it. The second use the pattern recognition techniques; there is only one video camera and some target markers geometric knowledge.

The fact that the inside-out method gives more accurate results and a better orientation resolution than the outside-in method should be noticed.

Registration

The registration is one recurrent problem in such system. In order to make the visualisation of the real object and its virtual change realistic, an accurate registration is required.

Two types of errors can be encountered: the static and the dynamic ones [9].

The static errors are due to the optical distortion, errors of the tracking system and differences between models or material specifications and real material physical properties. This kind of errors is perceived even if the user does not move.

The dynamic errors are due to the processing time lag that is the delay between measurements made by the tracking system and the display of the virtual entity. In fact, there are due to all processing time devices and systems. Different ways have been explored to reduce this dynamic error, by reducing the system lag or the perceptible delay [11], by making location prediction [11], by image matching.

To make a good registration the tracking system must perform good locations.

To resolve the static errors a calibration have to be made, and as many system, ours will use Kalman filtering to reduce dynamic errors.

SYSTEM ISSUES

Hardware issues: Optical see-through Head Mounted Display

For the moment this technology is not mature enough. As we saw upper, there are two main devices, the Sony

Glasstron and the Nomad of Microvision. Their display length on the field of vision can be considered as small.

Devices	Resolution	Display length
Sony Glasstron	800 x 225	
Microvision HMD	800 x 600	23° x 17°

There are other display devices on market, based on the head mounted display concept, like video-see through HMD and some video screens that can be clipped on glasses. However, the first one does not allow users to see their close environment directly, and the second type proposes a small display of a computer screen, PDA screen, etc.

Real object location issue

The real part is on a There table. is visualisation problem (Cf. figure 9) different changes are made in different locations of the real part (one of the bottom and another at the top for instance).

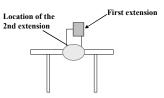


Figure 9 - problem visualisation

The real object location needs a thorough thought in order to permit users the visualisation of different changes.

Gestures recognition issues

Gestures recognition using visual approach, in particular without the use of markers, is not a mature technology. Some systems work in real time but in very specific conditions, such as with a uniform background, a small vocabulary, etc. Even if a system that recognises gestures, with a visual approach without markers, in real time, in a normal environment would add a tremendous value to general MR systems, it remains, for the moment, a research perspective.

Moreover, as people know, in the human-human communication, the use of speech, gestures, and facial expressions contributes to the information exchange. In particular during argumentative phases or solution negotiation activities, the frequency of gestures and facial expressions significantly increases. The case study presented in this paper deals with one specific situation where people will try to negotiate integrating several different points of view. So, making the difference between social communication gestures with actual gestures performed to interact with the system is primordial in our case. The system must have the possibility to distinguish these two types of gesture, in order to avoid false interpretations.

Virtual modifications

Object that people can bring in a meeting room are generally medium-sized. Moreover, changes made on a part during the review sessions are not consistent; the general object shape is not called into question as modifications are localised and concern only small areas. Therefore, to keep the visualisation as realistic as possible, the registration of the virtual change has to be very accurate.

Part modification is not necessary an extension of the real object, it could be a suppression of a little part area. To make an extension the system proposes to the user to mould a classic geometric form and after to register it on the real object. But to make small area suppression, it is more difficult to make it realistic, the representation is more elaborated.

The system has the three-dimension model of the real object. This will help us to make this kind of modification as realistic as possible.

Our first solution is to use a dark shape to simulate the suppression area. But the realistic visualisation depends too much on the user's location (Cf. Figure 10).

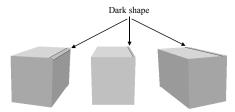


Figure 10 - different modification views

Occlusions issues

Occlusion is the typical issue of this type of MR system.

All actors are situated in a close room area. In our study case, the design review is made around a table; so all users move and perform their task, in a confined place. In this context, occlusions will be frequent.

Solutions must be found in the room layout, in cameras positioning and in tracking methods.

CONCLUSION

This paper deals with information sharing in our MR system for design engineering, the modalities used, the

tracking system and issues met. All these bases will allow us to build a prototype.

REFERENCES

- 1. Rauterberg, M., From Gesture to Action: Natural User Interfaces. Eindhoven University of technology; Mens-Machine Interactie, Diesrede 1999, 15-25.
- 2. Fjeld, M., Bichsel, M. and Rauterberg, M.: BUILD-IT: Intuitive plan layout mediated by natural interaction. *Arbete Människa Miljö & Nordisk Ergonomi (work, human being, environment), 1/99*, 49-56.
- 3. Romell, O., Arthur (augmented Round Table for Architecture and Urban Planning) An augmented Reality, Task Oriented Collaboration Environment. *AVR II and CONVR 2001 Conferences at Chalmers*, Gothenburg, Sweden, October 4-5, 2001, 29-37.
- 4. Grasset., R., and Gascuel, J.D.: MARE: Multiuser Augmented Reality Environment on table set-up.
- 5. http://www.immersion.com/3d/products/cyber_touch.
- 6. http://www.immersion.com/
- 7. Ying Wu, Thomas S. Huang, "Vision-Based Gesture Recognition: A Review", A. Braffort et al. (Ed.) Lecture Notes in Artificial Intelligence 1739, Gesture-Based Communication in Human-Computer Interaction (International Gesture Workshop, GW'99, Gif-sur-Yvette, France, March 1999).
- 8. Ying Wu, "Vision and Learning for Intelligent Human-Computer Interaction", *Ph.D. Dissertation*, 2001, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign.
- Rolland, P.J., Baillot, Y., and Goon A., A survey of tracking technology for virtual environments. Center for research and education in optics lasers (CREOL), university of central Florida, Orlando FL 32816, 1999.
- 10. Azuma, R., A survey of augmented reality. *In presence: teleoperators and virtual environment* 6, 4, August 1997, pp.355-385
- 11. Kijima, R., and Ojika, T., Reflex HMD to Compensate Lag and Correction of Derivative Deformation, *IEEE Virtual Reality Conference*, Orlando, Florida, March 24 28, 2002, pp.172-182.