A View on the Design of Usable Museum Appliances

Piero Mussio[(*)] *and Augusto Celentano*[(#)]

[(*)] Dipartimento di Elettronica per l'Automazione, Università di Brescia, Italy
[(#)] Dipartimento di Informatica, Università Ca' Foscari, Venezia, Italy
E-mail: mussio@ing.unibs.it, auce@dsi.unive.it

ABSTRACT
This paper poses the problem of designing Museum Appliances (MA), which are usable, i.e. easy to learn and easy to use. A recently proposed model of multimedia Human Computer Interaction is used to explore the dimensions of the MA usability, identifying the features which characterize this type of appliances, and to outline the an approach to usable MA implementation. The reasoning is bottom up, in that it starts from the study of some existing MA within the frame offered by the model, to derive the features which characterize an MA and the usability requests to be satisfied.

INTRODUCTION

Museum Appliances (MA) are complex systems designed to support humans in their interactions with a virtual or real museum. An MA allows its users to navigate in the real or virtual museum space, presenting them data which document the current state of the interaction and how the interaction can be continued. However these data can be correctly interpreted only by users who recognize their nature and know the rules for their interpretation, i.e. by users who belong to a certain culture. But MA users have different cultures, skills and dexterity levels and interact with the system in different contexts to achieve different tasks.

The problem arises of designing *usable* MAs, that is, easy to learn and use [8]. This paper addresses this problem by capitalizing on a recently proposed model of Human Computer Interaction [2] and on several practical experiences[4]. Users–MA interaction is studied in the frame of this model identifying a set of communicational and operational features an MA has to posses to be usable. Indications are derived on the design of MAs which enjoy those features which are culture dependent. An architecture is thereafter presented which is able to customize the ways in which MA interacts with its users adapting itself to their culture.

MUSEUM APPLIANCE

Museum Appliances are software–hardware multimodal interactive systems designed to support users in their interactions with a virtual or real museum. They can be classified in several categories according to their primary purpose – e.g., supporting user orientation and movement, providing introductory or shallow information, providing specialized information for professional users, and so on –, to the complexity of their operation, to the completeness of information they can provide and to their mobility. Not surprisingly, the more they grow in complexity, the wider is the information they can furnish. We could adopt a simple classification scheme centered on three basic kinds of devices:

- Hand held devices, primarily audio

devices, whose purpose is to provide a shallow information for the novice user, and a help in following guided tours.

- Kiosks, i.e. multimedia devices able to give a variety of information for users of different skills. An extension if kiosks are CD-ROM based guides and electronic books. Both are based on the same technology, but they serve different purposes.
- World Wide Web sites. They could be considered extensions or technological variations of CD-ROM, which can provide dynamic and open information Many museums have set up Web sites that help the user to set up a tour, or provide practical information, or serve as an anticipation of what the visitor could see at the museum.

We base our discussion mainly on visual interfaces typical of CD-ROM and Web based appliances and address issues related to information access, navigation and orientation.

MA provide three main supports to their users, which allow them to perform their tasks without loosing themselves in the information space:

- maps, i.e. orientation devices that help the user to understand where he/she is, from where he/she came, and where can go further;
- guided tours, i.e. orientation devices that help the user in stepping through the information in a regular and motivated way, according to a knowledge goal;
- multisensory devices that supplement knowledge about the user movement (physical or conceptual) with easy to perceive *stimuli* that integrate rather than replace the information he/she is accessing at any moment.

The MA supports its users by sending them multimodal messages: these messages can only be understood if the user recognizes them as messages and is able to decode them. To make this point clear, let us frame the user–MA interaction in the Pictorial Computing Laboratory (PCL) model of WIMP Human Computer Interface [3].

Without loss of generality, we first consider a WIMP MA, in which the interaction is based on the exchange of images between two participants, namely the human user and the MA itself. Both human and MA interpret every event occurring during interaction with reference to the whole content presented in the MA screen display. The content is formed by text, graphs, pictures, icons, etc., therefore representing a multimedia message, and materializes and conveys the meaning intended by the sender. The message must be interpreted – i.e., associated with a possible different meaning – by the receiver. The screen is the communication surface between human and MA, thus acting as a bi-directional channel in the transmission of messages. The receiver interprets the message based on cognitive criteria which could be described as more or less formal rules (or conventions) which come from experience, habit, cultural background, etc.

Figure 1. A list of paintings from WebMuseum

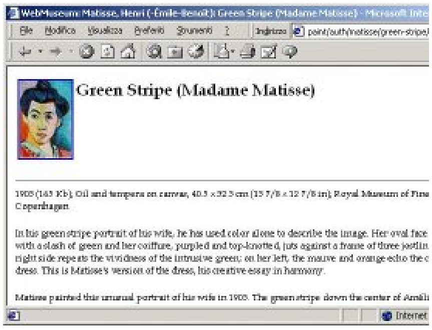

Figure 2. A painting description

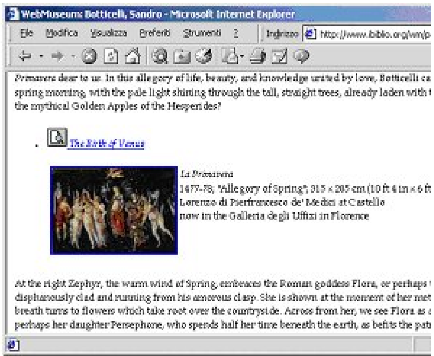

Figure 3. A comment on Botticelli's paintings

INTERPRETATION OF MUSEUM DOCUMENTS

Figures 1–3, taken from the WebMuseum [9], help to understand this process, showing how users can interpret museum documents based on visual properties and layout.

Figure 1 is a list of references to paintings, each of which is introduced by a small image and a technical description. The image is an index and has a descriptive function: it gives the user a glimpse of the artwork. The description is perceived as an introductory, objective, technical one due to its limited size and the presence of recurring text patterns (Title in italics and painting size being the most immediately perceivable ones). A second type of indexes is used**:** a symbolic icon (a magnifying glass) coupled with a string of characters suggesting that a more detailed description of a painting whose title is represented by the string can be reached, as in Figure 2, which shows a separate document with a simple visual layout, suggesting no further details. The absence of visual cues in the text asides the paintings in Figure 1 suggests that only some artworks are described more thoroughly.

In Figure 2 the small size of the image suggests that it is an index, as in Figure 1, while the blue frame which surrounds it suggest to whoever is a Web surfer that a larger image should be accessible by clicking over it, blue being the traditional color for links in WWW documents. In the large-sized image standing as a separate document it is evident that no other information is associated to it.

Some structures in each image on the screen act as cornerstones, i.e. they help the user in navigating through the virtual space: typically the image title, which identify the overall virtual space, i.e. the painter Matisse in the Web Museum, and the situation, i.e. the *Green Stripe* painting. The user is able to interpret the purpose of each document after a brief glance at its layout. A brief tour in the WebMuseum site shows that these patterns are recurring, even if sometimes deviations from this style can be found. An example of a different pattern is in Figure 3. This document is not a simple list of paintings: a criticism on Botticelli's artworks discusses two paintings, the *Birth of Venus* and *The Spring*. The former is described in a separate document (hence the icon), while the latter is described inline

(hence the small painting image). The choice of organizing the discussion in this way is an author's choice and is not motivated in the document.
The interaction process with an MA can thus be modeled according to the PCL model, which was however derived from the study of critical interaction processes, in which every mistake has a defined cost. This is not the case of MA interaction. This difference in the context of application will lead to different criteria for their design, even if the same model of interaction is adopted. The interaction process is in both cases modeled as the generation of a sequence of images appearing on the screen at successive instants of time t_1, ..., t_n, each one interpreted both by the human and by the MA. Each image derives from the transformation of the previous one according to user's actions and MA's computational activities.

Human and MA communicate by materializing and interpreting a sequence of images. Humans interpret and materialize the images using human natural cognitive criteria, the MA using the criteria programmed in it by the system designers and programmers. Two semantics are always implicitly defined in any interaction: one internal to the MA, in which each image is associated with a computational meaning, as defined by the designer and implemented in the MA; and one proper to the user performing the task, depending on his/her role in the task, as well as on his/her culture, experience and skill.
Ambiguous or equivocal situations may arise because of these two interpretations: *Ambiguity* arises when one of the two communicants – the human or the computer – associates two different meanings with a message, and flips from one meaning to the other during a same reasoning process. *Misunderstanding* arises when the two communicants associate two different meanings with a same message. Each communicant has no ambiguity in what he/she is doing, but different cultures or situational reasons determine different interpretation. Conflicts may arise between the two communicants, which make the user unhappy.

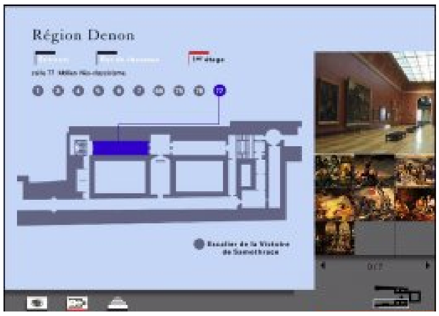

Figure 4. A map from *Le Louvre*

As an example let us consider guided tours and maps. The goal of a guided tour is defined by an author who could rely on a culture different from a visitor, therefore designing a tour not meaningful in the visitor cultural context, or difficult to understand. Maps should be more objective because they document a real or realistic layout, but ambiguity can arise also in this context. Figure 4 shows a floorplan of a region of the Louvre Museum, excerpt from the CD-ROM *Le Musée du Louvre* [6], with a room selected. The room is highlighted on the map in blue, and pictured on the upper right side of the screen. On the lower right side a mosaic of the paintings exhibited in the room and contained in the CD is shown. The floorplan and the mosaic are clearly maps. The room image is not perceived as a map in this context: it is only a preview of the real museum room. A click on this image shows the screen of

Figure 5 (frames are visible only when the cursor moves over a painting, so Figure 5 is actually an artifact obtained by several superimposed screen shots). Differently from the preview image this screen is a clickable map used as an intermediate step to reach the painting documents.

Figure 5: A room acting as an active map

As observed in [5], users can achieve their tasks if they associate with each structure in the image and with the whole image a meaning similar to the one associated by the MA, i.e., if an *adequate communication* is reached.

REQUIREMENTS ON DESIGN

According to the PCL model, an interaction process is based on two interpretation and two materialization processes. From this point of view, the goal of a successful design is to bring the system semantics to reflect the user's one, so that both the user and the system materializes messages which are properly understood by the user and adequately managed by the MA. Moreover users should be in the condition to perform actions, familiar to them, in a way that is correctly perceived by the system.

The MA designer has to take into account both the communication aspects of the interaction, regarding how the exchanged messages are interpreted, and the operational aspects which regards how the messages are materialized. At the communication level design must consider:

- variety: multiple facets of the same information (images, text, sounds, but also descriptions, criticism, comments) able to convey complementary meaning to a user in order to improve his/her knowledge about that information item;
- customization: a form of personalization targeted to the specific user needs and habits beyond the ones common to a wider cultural or geographic environment;
- deepening: the ability of navigating into an information item to discover new more specific information according to a progressive disclosure principle;
- accessibility: the ability of knowing what further information exists, and how it can be found.

At the operation level design must consider:

- navigation: the ability of moving among different information items effectively, i.e., without missing information due to the difficulty of understanding how to move;
- orientation: the ability of knowing at each stage of the visit where information is, what information has already been examined, what new information can be accessed;
- personalization: the ability of accessing the information in a way consistent with the cultural habit of the user;
- retrieval: the ability of finding an information item without knowing the path to reach him, or the knowledge that an information item is definitely missing.

A fundamental distinction between cultural heritage appliances and industrial systems appliances is in the use of metaphors in a user interface. Industrial systems make large use of metaphoric information for representing information and data about the real world. Metaphors must be easy to understand, therefore simple and related to a common understanding. Cultural heritage systems render real world information in such a way that it can be recognized as a representation of the real world, mediated only by the technology that displays it to the user. Therefore there is no need for common understanding or adoption of standard conventions, unless they are targeted to giving the user a sense of realism. For example, the use of the same scale for representing a set of paintings gives the user the knowledge about the relative sizes of the real worlds works; the use of a same colour resolution (for paintings) or lightning (for sculptures) makes easier for the user to compare different works. Metaphors are confined only to the operational part of the interface, and not to the informative part.

On the whole, an MA designer has to pursue user satisfaction, as an industrial appliance designer do, but in a different context: Museum Appliances and Industrial Appliances are designed with different goals, a mistake of the user as well as wrong behaviors have different costs and their success must be judged according to different usability metrics. Moreover MA users may belong to a many different cultures, which all are valid interpretation tools of the Museum content, while in an industrial environment the users' culture variability may be considered more restricted.

However at a high level of abstraction, the guidelines to correct design are similar: (1) The designer has to develop a *usable* MA: the system is easy to learn, to use, to navigate, efficiently and effectively. However, easy to learn bears now a peculiar meaning, due to the fact that an MA may be designed to be used just once by a user – think of a hand held device, which supports museum visitors. (2) The designer has to develop a *viable* MA: no human action result into a non-meaningful situation or in a system crash. (3) The designer has to develop a *deterministic* MA: in every situation an action has to produce only one (predictable) result.

In the interaction, no message can be ambiguous: the human (or the computer) never assigns to a message different meanings during the same interaction process and the communication must be adequate: i.e. both human and computer have to assign the same meaning to each message. These two features are difficult to be achieved, because of the variety of cultures to which MA users may belong to, and also to the difficulty to obtain a user profile. Last, the user navigating in the museum space (real + virtual) must always understand where he/she is and never gets lost in the virtual space.

FORMALIZING THE INTERACTION PROCESS

Let us now examine the interaction process in more details, following the PCL model. The images on the screen represents the state of the interaction, and their interpretation determines the next action of the human and the next computation of the MA. Humans interpret the images on the screen by recognizing *characteristic structures* (*cs* for short), i.e. sets of image pixels that users recognize as functional or perceptual units. The *cs* recognition results into the association of a meaning with a structure. Humans express the meaning attributed to the *cs* by a verbal description. In an image a human may recognize several *cs*: combining their meanings, humans derive the meaning of the whole image on the screen. Hence we can say that *cs* play different *roles* in the user understanding of the image.

For example, in Figure 1 the two paintings are recognized as *cs*. In this situation they play both the role of *indexes*, indicating to the user that the painting is present in the archive and that of *links*: the painting can be seen in full screen mode by clicking on it. A different *cs* is constituted by the rectangle with superimposed a file symbol and the magnifying glass icon. This last *cs* can be seen as a composed one, and its meaning can be derived only combining the meanings of its component *cs*. Several other *cs* can be recognized: every users, who knows Latin alphabet, is able to recognize the single letters appearing on the screen, while a user who also understands English recognizes the words and sentences, which therefore become complex *cs* for him/her. For this user the title in the top bar plays the role of a cornerstone, which help him/her in orienting in the virtual space. A user who uses a different alphabetic system can suspect that those structures may represent letters, while a user accustomed to different writing techniques, e.g. pictorial ones, may not even recognize the role of these *cs*. On the whole, the screen shown in Figure 1 can be properly understood by a user who knows English language and is familiar with the aspect of traditional desktop interfaces icons and web surfing.

On the other hand, the MA associates graphical entities on the screen with computational constructs. Each computational construct – here denoted by *u* – represents the meaning associated with a graphical entity on the screen, as intended by the system designer. It is exactly this association that makes the MA able to interpret the captured user actions such as clicking on a button, with respect to the image on the screen, possibly firing computational activities whose results are materialized on the screen, via creation, deletion, or modification of *cs*. The association between a *cs* and the corresponding *u* is called a *characteristic pattern* (*cp*).

These concepts can be formalized starting from the definition of a characteristic structure *cs* as a set of pixels perceivable on the screen and of a characteristic patters *cp* as a triple which specifies how the set of pixels (*cs*) is linked to a program *u* which describes its meaning.

In an image *i*, several *cs* can be identified. These *cs* became *cp* when a program *u* is associated to them. For example, in Figure 1, the two painting images are *cs*, whose description contains links to a program which changes the image on the screen, presenting the painting on a full screen. Moreover, *i* as a whole can be associated with a program *d* synthesizing the overall properties and the global meaning of the image *i*. We call *visual sentence* (*vs*), a triple whose elements are the image *i*, the program *d* and a specification of the relations among the *cs* in *i* and the programs which describe their meanings. A set of *vs* is a *Visual Language* (VL).

A formal theory that provides a finite definition of VLs by a special family of rewriting systems, the *Visual Conditional Attributed Rewriting Systems* (*vCARW*), has been described in [1, 2]. In this theory an interactive process can be specified by the possibly infinite set of all the sequences of *vs* that, starting from an initial *vs*, vs_0, are determined by the sequences of user actions and system computations in the process. Each sequence in the set describes a specific user-computer interaction session. The set of all sequences of *vs* that can be generated from vs_0 constitutes the *Interaction Visual Language (IVL)*.

On the basis of this definition, adequate communication occurs when in each *vs* of IVL, the graphical entities recognized by the MA are the *cs* recognized by humans, and the behavior of associated computational structures reflects the behavior expected by the users.

A STRATEGY TOWARDS REQUIREMENTS SATISFACTION

The recognition of the existence of two semantics is the starting point for reaching an adequate communication between human and computer. Indeed, adequate communication occurs when the human and the computer associate a similar meaning to a same message (or part of it) [5]. In order to identify the users' semantics, the PCL recommends to exploit the users' notation as the kernel of the definition of the IVL through which human and computer communicate. Users' notation embeds context, task and procedural knowledge possessed by the users. It embeds the knowledge explicitly and implicitly: explicitly in symbols and rules used for constructing the users' documents; implicitly, because the shape and the spatial arrangements of symbols allow users expert in the application domain to recognize structures and relations meaningful for their tasks, but too vaguely defined to be made explicit in a declarative format.

The adoption of users notations as the kernel of the IVL is a fundamental step to make the interaction process understandable and checkable by the users: users may justify results expressed in their notations on the basis of their experience and not on the basis of an algorithmic explanation. In this way, the adoption of user notation facilitates the reaching of the closeness of mapping between the real world, in which users operate, and the virtual world which supports their work. However, as pointed out by Mayhew [7], the mere adoption of this notation may bring to under-use the system. In fact, the users' notation has the advantage of being completely familiar to the user, but also the disadvantage of having been defined without taking into account the existence of computing systems. The PCL approach therefore proposes to *augment* and *adapt* the original notation to fully exploit the computing capability of the interactive systems. In the augmented language, symbols are able to show their state, for example, assuming a color to show that they have been selected, and can be associated to a specific functionality to favour the interaction with the user. For example, traditional supports exist also in real world museums (e.g., maps, arrows, colored pathways) and many studies exist on museum architecture.

Museums Appliances augment traditional tools, because they also support users in understanding and checking the information, mainly helping to understand the overall information organization. While in an industrial application the user interface for operating a complex apparatus must make evident what the industrial system does at any stage of operation, in a cultural heritage application the interface must make evident to the user what information he/she can find, in which form, and at which extent. Since MA users have different cultures, skills and dexterity levels and interact with the system in different contexts to achieve different tasks, different styles of augmentation and adaptation should be adopted to develop MAs for different users communities. What we are studying is to allow MAs to be adapted by outer actions or by self-evolution to the user according to some user profile.

One possible approach, which we are exploring, recognizes the importance of the role played by the *cs* in the user interpretation. *Cs* may have the role of cornerstones, which help the users in orienting himself in the virtual space; may indicate the possibility of executing an action; may convey information on the subject of investigation. *Cs* are distributed in the screen space according to their role, and often their appearance depends on it (an alarm or warning in west Europe is a red flag). Shape and position of a *cs* help the user in recognizing and in understanding it. Unluckily, these features depend on the user culture: users who read a text from right to left expect the title cornerstone on the right side of the screen. To take into account these features, the data related to a Museum Appliance are coded in an intermediate form, typically using an extension of some tagged language. Tag types are introduced which define the role that each bunch of data plays in the communication. For example, they specify if a text acts as a cornerstone, and should orient the user in the navigation within the Museum space (real + virtual); a tag related to an image specifies if the role of the image is to be an index, or a link, or both; in a sequence of images, tags specify their time dependency.

This intermediate form can thereafter be translated by an instance of MA, suited to a given user profile. Museal data are presented to the user according to his/her habits and needs. In other words, the MA not only translates the text, but also interprets the tags to adapt the layout, the symbols and the operational facilities to the specific user culture, skills and dexterity. Each MA receives the description of the museal space coded in the tagged language and interpret it for its users. The MA also accepts the users input in the user notation and interpret it to match with the Museum description. This organization can be achieved by designing the MA as a bi-directional translator among the user notations and the tagged language.

CONCLUSION

This paper stems from some practical experiences from which a classification of MA is derived as well as some usability problems emerged. MA are take-up-and-walk tools, which need to be understood at once by their users, who however have different goals and belong to different cultures. To face this problem, our proposal is to design MAs as bidirectional translators customized to a specific community of users, that is, users which have common goals and belong to a same culture. The museal space is described and maintained only once in a tagged language. Each MA customized for a specific user community translates the data of the museal space into the user notation and interprets the input from the users, expressed in their notation, with respect to the intermediate form.

REFERENCES

1. P. Bottoni, S.-K. Chang, M.F. Costabile, S. Levialdi, P. Mussio. Dynamic Visual Languages, *Proc. IEEE Symposium Visual Languages'98*, 14-21, 1998.

2. P. Bottoni, M.F. Costabile, P. Mussio. Specification and Dialogue Control of Visual Interaction through Visual Rewriting Systems. *ACM TOPLAS*, Vol. 21, N. 6, pp. 1077-1136, 1999.

3. P. Bottoni, M.F. Costabile, S. Levialdi, P. Mussio. A Visual Approach to HCI. *ACM SIGCHI Bulletin*, *Vol.28, No.3*, July 1996, pp.50-55.

   http://www.acm.org/sigchi/bulletin/1996.3/levialdi.html.

4. A. Celentano, F. Pittarello. Multi-sensory guided tours for cultural heritage: The Palazzo Grassi experience.
   Technical Report, Ca' Foscari University, Venezia, Italy

5. S.K. Chang, P. Mussio. Customized Visual Language Design. *International Conference on Software Engineering and Knowledge* Engineering - *SEKE'96*, pp. 553-562, 1996.

6. *Le Musée du Louvre, les collections*. Montparnasse Multimedia & Réunion des Museaux Nationaux, 1997.

7. D.J. Mayhew. *Principles and Guidelines in Software User Interface Design*. Prentice Hall, Englewood Cliff, 1992

8. Nielsen, J. (1993). *Usability Engineering*. Academic Press, 1993.

9. WebMuseum is available at several URLs: *http://www.oir.ucf.edu/wm/*, *www.ibiblio.org/wm/*; other URLs are referenced by these sites.