Adaptive, intelligent presentation of information for the museum visitor in PEACH

**Oliviero Stock · Massimo Zancanaro · Paolo Busetta ·
Charles Callaway · Antonio Krüger · Michael Kruppa ·
Tsvi Kuflik · Elena Not · Cesare Rocchi**

Abstract The study of intelligent user interfaces and user modeling and adaptation is well suited for augmenting educational visits to museums. We have defined a novel integrated framework for museum visits and claim that such a framework is essential in such a vast domain that inherently implies complex interactivity. We found that it requires a significant investment in software and hardware infrastructure, design and implementation of intelligent interfaces, and a systematic and iterative evaluation of the design and functionality of user interfaces, involving actual visitors at every stage. We defined and built a suite of interac-tive and user-adaptive technologies for museum visitors, which was then evaluated at the Buonconsiglio Castle in Trento, Italy: (1) animated agents that help motivate visitors and focus their attention when necessary, (2) automatically generated, adaptive video documen-taries on mobile devices, and (3) automatically generated post-visit summaries that reflect the individual interests of visitors as determined by their behavior and choices during their visit. These components are supported by underlying user modeling and inference mecha-nisms that allow for adaptivity and personalization. Novel software infrastructure allows for agent connectivity and fusion of multiple positioning data streams in the museum space. We conducted several experiments, focusing on various aspects of PEACH. In one, conducted

with 110 visitors, we found evidence that even older users are comfortable interacting with a major component of the system.

1 Introduction

The PEACH project (Personal Experience with Active Cultural Heritage) has been a large endeavor concerned with developing a novel integrated framework for museum visits. We have taken up the challenge of developing various intelligent technologies and assembling them into a coherent picture to support and improve visitor interaction and satisfaction.

The main idea has been to produce a multifaceted system that accompanies the visitor and augments her overall museum experience. Our guiding principles are that the technology that accomplishes this must not be obtrusive, that nothing should come between the visitor and the exhibit and that the emotion of being there in front of the real thing ought to be fully preserved and if possible enhanced even further. This paper is intended to serve as a reference for some of the main achievements of the PEACH project related to adapting intelligent user interfaces in museums for individual experiences.

Before giving an overview of the elements of the PEACH project, let us examine the characteristics of the museum domain that motivate the widespread interest in the intelligent user interfaces sector.

Setting: Museum applications include both mobile devices and static fixtures like kiosks, as well as interactions with both the exhibit environment and its associated information space. The wide variety of settings opens up many challenges to current research, including mobile interfaces, pervasive computing, and ethno-methodological studies.

Interaction quality: interaction is over a prolonged period of time in a coherent manner compared with other domains amenable to adaptation such as web browsing. Integration of multiple sources of data, explanation and comments must add to the main focus of attention: the exhibit itself. The combination of an overall story with insight on its details is needed along with ways to keep attention levels high. Research on all kinds of communication technologies, multimodal presentations, user modeling and user adaptive techniques can contribute here, as well as different methodologies of user studies.

Audience: Spanning individuals and groups, adults and children; audience is also relevant for other important constituencies, such as old or impaired people. Specific devices and multimodal interfaces must be separately designed and studied for different audiences.

Visitor motivation: In a museum, people are motivated to gather more information, to have fun and learn. In fact, they come to the museum for the combination of emotional experiences and learning. The study of the role of emotion in cognitive processes has recently opened the way to investigations aimed at developing technology that gives a role to emotion in interaction and cognition, and the museum is a privileged setting for such activities.

Social value: the underlying goal of technology in the museum is to help promote cultural heritage and a positive attitude towards learning and appreciating culture; it is not hard to find a setting of local interest, making constructive experimentation attractive in terms of social impact.

Design and implementation challenges: Prototype construction is essential, with all the technological and methodological challenges this implies. User studies *in situ* are both another challenge and opportunity, and multiple methodologies are needed to obtain meaningful

results. Such studies must be performed iteratively and repeatedly whenever prototypes are updated.

Commercial potential: industrial turnover for mobile electronic guides is not yet enormous (US $67 million in 2004, according to Espro, a major player in the market), but it is widely believed that it will grow substantially in the coming years. New technologies will also impose an evolution on the business models prevailing in this market sector.

Risk: museum visits are still a low risk setting. It is not "safety critical", and suboptimal solutions are often acceptable; there is ample space for advancing the current state of the art. Though success can be reasonably defined—and may not be always attained—dramatic failure is something that is simply not to be feared in most projects.

These motivations and opportunities explain the large number of projects that have been carried out in the museum domain, and below in Sect. 1.1 we describe some of the most important. Various projects have introduced technology for mobile presentations triggered by the position of the visitor in the physical space. Such technology typically takes advantage of a localization system (for instance based on devices that generate an infrared signal from fixed positions, based on triangulation through emitters/receivers of wireless digital signals, or on very sensitive GPS systems, which recently have achieved enough sensitivity to work inside buildings). The visitor typically carries a small portable device (for example a PDA or a simple headset), and receives information relevant to items at their particular location.

In the PEACH project we have considered the scenario of technology-assisted museum visits with a broad perspective. We have seamlessly integrated both mobile and stationary components into the system so that one can for instance stop at a large screen and obtain a presentation appropriate for that device before moving on with presentations on a mobile device that continue with or otherwise take into account the prior interaction. The system is initialized with a user profile and then, in the course of the visit, adapts to the behavior of the visitor, proposing personalized, context-dependent presentations. Presentations themselves are multimodal, and in particular we have developed technology that combines language presentations and small, tailored, visual documentaries, meant to provide a coherent network of support for the visitor, for instance when guiding the visitor's gaze toward particular details of an exhibit.

Another component of the system aids in identifying the focus of attention of the visitor when in front of a large bi-dimensional space, like a fresco. The artificial vision-based technology pinpoints what region the visitor is currently looking at so that relevant presentations about the details of that area in the focus of the visitor's attention can be provided.

In the course of a visit, the system capitalizes on each individual visitor's feedback whenever possible to guarantee appropriate presentations for their interests and taste. The system "observes" the visitor's behavior, and the elements comprising that behavior are interpreted as implicit input. Explicit input, on the other hand, is gathered in a very simple and intuitive way: feedback on the part of the user is limited to simple communication of liking or disliking of the current piece of the presentation, with consequences for subsequent presentations. And at the end of the visit, a personalized written report that summarizes the key aspects of the overall visit experience is automatically produced for the visitor to take home with her or to get as an electronic diary and an entry point to their subsequent cultural experiences.

A variety of topics related to the PEACH scenario have been the occasion for development of substantial research:

- *Architectural infrastructure*. A first step was to define a software infrastructure that can accommodate the many different fixed and mobile devices and subsystems. Components of the system need to both communicate directly and to "be aware" of what is happening

in the communication between other components. A new agent-based infrastructure was designed specifically for ambient intelligence applications.

- *User Modeling*. Supporting personalization is a particularly challenging task in the museum environment where "non-intrusiveness" is fundamental. User modeling has to take advantage of the observation of explicit visitor choices and requires integration of various UM techniques. Our distributed architecture was essential for introducing an element of competition and collaboration among these different techniques. Two prototypes in particular use UM data—dynamic presentation and report generation.
- *Personalized multimedia presentations*. Significant effort was put into research in automatic documentaries, which can be produced dynamically and adaptively from scratch using a detailed user model, a corpus of static pictures and content and rhetorical structure that are automatically combined into the text of the presentation.
- *Seamless ubiquitous presentations*. A fundamental issue has been how to give coherence when stitching together a whole set of presentation components. We have come up with various techniques allowing the visitor to perceive that synthesized coherence. For instance: (a) when the visitor is moving from a presentation provided via a stationary device (a large screen) to a mobile device (a PDA), an animated character "jumps" from the former to the latter and continues the presentation, with the specific point of view its appearance embodies; (b) the presentation text makes references and comparisons to exhibits this specific visitor has seen; and (c) a final report becomes a compact souvenir of the specific things seen and presented and of the visitor's apparent points of major interest.
- *Report generation*. Specific challenges in text generation, concerned with how to produce a good quality summary report of a visit experience were another subject of substantial research, along with integrating the report generator with the user modeling component to ensure that personalization is not only present, but is evident to the visitor.
- *User studies for interaction design and evaluation*. Another important object of study was the user interface: how best to obtain feedback from the user, which was simple, intuitive, effective and meaningful. These challenges have required design and evaluation user studies that have led to some significant advancement in the state of the art, for instance our empirical discovery that video-clips that accompany spoken text on a PDA do not require more attention shifts from and to the real object than do still images.

PEACH has resulted in the creation of a multifaceted prototype that we have experimented with at a historic-artistic Renaissance site at the BuonConsiglio Castle in Trento, and, in part at an industrial archeology site, the IronWorks at the Voelklinger Huette museum in Saarbruecken. Prototype evaluations with users were carried out on the former site.

1.1 State of the art of intelligent mobile guides

The museum visit is a complex experience that encompasses multiple cognitive and emotional aspects. Visitors have different ways of elaborating background and new knowledge, and they have different interests and preferences. Designing interactive technologies to support such an experience requires effort in a number of aspects: the graphical interface and its usability, the adaptation mechanism and its effectiveness, and the overall satisfaction of the visitor. Moreover, technology designers need to consider the intrusiveness of the devices they propose for supporting visits.

Applications of wearable or mobile computers initially focused on traditional tasks in the 'office domain', such as accessing email, checking a calendar, etc. In general, the tendency in the beginning was the re-implementation of 'desktop tasks' on mobile devices. More recent studies have attempted to enlarge the application domain of mobile devices, by introduc-

ing new functionalities such as position and context-awareness (Want et al. 1995; Cheverst et al. 1999), information sharing (Aoki et al. 2002), and adaptivity as discussed in this paper. Studies involving mobile guides range from architectural issues (communication protocols, frameworks, portability, etc.) to specific functionalities to interface evaluations (statistical or ethnographical studies on acceptability, effectiveness, and usability). Here we review related work concerned with these issues, including three projects that have had a major impact on research on the possible use of mobile devices in various contexts: CyberGuide, HIPS and Guide.

CyberGuide has been one of the most influential works on mobile services (Long et al. 1996; Abowd et al. 1997). It describes a set of prototypes of a mobile, hand-held context-aware tour guide. The authors report on indoor and outdoor prototypes, the specific issue of detection of users' position and orientation and discuss a set of issues that emerged during its development, namely hardware choice, communication media and map representation.

HIPS has been a project focused on hyper-interaction (Benelli et al. 1999); it evolved from the earlier HyperAudio experience (Not et al. 1998; post hoc considerations in Petrelli and Not 2005). Its novel contribution has been the overlapping of contextual and personalized information on top of the physical space. In this way the user experience is enriched, because both spatial and informational navigation are allowed at the same time. The system was showcased in the Museo Civico (Siena, Italy), where visitors could move around, seeking information and guidance through a hand-held device.

The Guide system (Cheverst et al. 2000) is another successful implementation to date. It is an outdoor navigation system which supports tourists when visiting the city of Lancaster. It combines mobile computing technologies and a wireless infrastructure to present information tailored to users' personal and contextual situations. Of particular value is its design, carried out in collaboration with experts, and the insights gained during the evaluations.

Additionally, several survey papers help assess the development of the field over the years along different perspectives. For instance Baus et al. (2005) present a survey of mobile guides, whether prototypes or commercial, whereas Raptis et al. (2005) attempt to classify current practices in the design of mobile guides for museum.

Systems emphasizing a particular innovative aspect worth recalling here include DiscoveryPoint, the Genoa Aquarium Guide and SottoVoce. Discovery Point is a small remote control-like device that allows users to hear short stories related to a work of art; it is in use at the Carnegie Museum of Art in Pittsburgh (Berkovich et al. 2003). The prototype is an audio system consisting of the physical device that the visitor holds and a special speaker, which delivers pinpointed audio and can be heard near the work of art.

Another example of a PDA application can be seen at Genoa's Costa Aquarium (Bellotti et al. 2002). The interface's basic element is a multimedia card that corresponds to each presentation subject, such as a particular fish or a fish tank containing several fish species. Each multimedia card provides users with content and touch-screen buttons that enable them to control content presentation and navigate between tanks.

Sottovoce (Aoki et al. 2002) has something different: it is designed to promote information sharing during the visit, as it supports shared playing of audio content between pairs of visitors, each using a PDA. Grinter et al. (2002) report on a study with the system conducted in a historic house and present interesting findings on the shared experience.

Another crucial issue discussed in the literature concerning context-aware and adaptive mobile guides is architecture definition. Dey et al. (2001) for example present an interesting attempt to define the notion of context and introduce a conceptual model and a set of methods to help drive the design of context-aware applications. Their proposal is a computational framework to speed up the implementation of context-aware applications.

Efstratiou et al. (2003) present a new platform to support the coordination of adaptive behavior of multiple applications. The coordination can be specified in terms of policies (they also present a formal language to define such policies), which allow one to describe adaptive actions on a system-wide level. For instance it is possible to define the level of intrusiveness of the system, e.g., whether to notify the user that the system has taken a given action.

The focus on the architecture for seamless services has been a topic for research in other domains. Krüger and colleagues focused on the growing need for connections between different interfaces to provide user-adapted seamless services in different situations. They worked on a route planning scenario, in which a user is supported by a central system and three different devices: a desktop route planner, an in-car navigation system and a PDA-based pedestrian navigation service (Krüger et al. 2004).

Evaluation studies provide insights into both design and reimplementation of prototypes. In addition to previously cited works (Grinter et al. 2002; Bellotti et al. 2002; Cheverst et al. 2000), the added value of systematic evaluation is shown in a related domain in (Bohnenberger et al. 2005). This work reports on improvements brought to a mobile shopping guide after an iterative evaluation cycle. In particular, the authors focus on usability issues, related to the PDA interface, on the task to be accomplished (buy items in the minimum possible time) and on the accuracy of the system in supporting the user.

1.2 Guides based on dynamic content and adaptivity and their relevance to PEACH

Context-aware personalization of museum content is one of the key aspects to help each visitor accommodate and interpret the visit according to his/her own pace and interests.

HIPS, mentioned above, had several versions, and work developed at IRST was the forerunner of PEACH. Its content personalization was based on user position, interaction with the PDA and the surrounding physical space, and experiments were carried out on a form of high level ("strategic") dynamic language generation of content through navigation of existing material.

Ilex (Oberlander et al. 1998) focused on "deep" generation of dynamic contents for the (virtual) museum. The objective of the project was the generation of descriptions of museum objects (labels), tuned according to a set of visitor specific features, such as: type of visit, interest, and changing knowledge through the visit. One of the novel aspects introduced by Ilex is the use of hypermedia, which allows the connection of labels through the use of links. The dynamic generation mechanism implemented in Ilex permitted new possibilities to 'experience' the museum: (i) labels, dynamically composed, can be tailored to the specific interests of a visitor; (ii) there is no need to follow a specific visit path: each visitor can freely move and receive information presented according to previously seen exhibits; (iii) if a visitor comes back again to visit an exhibit she gets new information, rather than the same as presented before; (iv) retracing the history allows the system to generate comparisons with other objects, enhancing the coherence of the visit. All these features are also implemented in the PEACH system. Although the systems might look similar, PEACH introduces a new medium besides text and images: the video. This novel aspect enriches the expressivity of the generated content but also widens the range of choices to be performed by the system in order to compose coherent presentations according to visitor specific parameters like history, interests and previous knowledge. Moreover the video, unlike texts and still images, is a time-based medium. This feature has an impact especially on the rendering and presentational form of the generated content, introducing the need for synchronization. We addressed this issue by designing a platform independent language, XScript (see Sect. 3.3.2), to describe the rendering of both static and time-based media.

A system very similar to Ilex is PEBA-II, which describes entities in a knowledge base via the dynamic generation of web pages according to user profiles, novice or experts. It is worth noting that both Ilex and PEBA-II focus primarily on linguistic realization of the text in a hypermedia framework, whereas our main concern in PEACH is the coherence of the whole set of presentations selected during the visit and the synchronization among the media. Yet, we consider Ilex and PEBA-II as 'platforms' for the implementation of adaptive systems, whereas PEACH is a full-featured system with built-in location detection, user modeling and adaptation mechanisms, working in a mobile setting.

In the field of adaptive hypermedia we can find systems that implement dynamic selection of material according to user profiles, a more limited feature of what is realized in PEACH. GUIDE, mentioned above, generates web-like presentations for tourists visiting a city. Papers about this project do not provide many details about its adaptation features. The main dynamic aspect appears to be the (proactive) addition of information related to nearby attractions according to the position of the visitor and her profile.

DEEPMAP (Malaka and Zipf 2000) is a system that suggests paths to explore through a city. The research is particularly focused on technological infrastructure (GIS, 2D and 3D rendering of maps) and the integration of such resources. Its adaptivity consists of the generation of suggested paths according to the preferences of visitors, before the beginning of the visit. In the PEACH system, unlike DEEPMAP, adaptation is an ongoing process throughout the visit, continuously taking into account the feedback of the user.

MASTROCARONTE is an in-car adaptation system which automatically generates suggestions about hotels and restaurants according to the situation (time, location) and the driver's profile (Gena and Torre 2004). It is similar to PEACH in that adaptation is a continuous process, which takes into account the changes of the context and the user interaction, though the adaptation strategies are not easily compared due to the differences of the application domain.

2 The Torre Aquila and interactions with PEACH

We now introduce a typical PEACH scenario involving our developed prototype technologies. When a visitor first arrives, she finds a stationary computer with a large plasma screen located at the entrance of the Torre Aquila, a defensive tower forming part of the Buonconsiglio Castle in Trento, Italy (Fig. 1a). The interior is painted with famous Fourteenth Century frescoes representing the Months of the Year with scenes of towns, nature, and daily activities that would have occurred during each month. The visitor is also given a Personal Digital Assistant (PDA) with infrared capability which helps the PEACH system determine general position and orientation based on coded emitters placed near exhibits, and which can also be used to interact with the exhibits themselves.

As the visitor's position at the entrance is detected, lifelike characters on the screen start a dialogue with the visitor. They introduce themselves, provide a general description of the tower, and each tries to persuade the visitor to choose it as a companion for the duration of the visit (Fig. 1b). When a character is selected it fades away from the stationary device and appears as an animated agent on the visitor's PDA (Fig. 1c). From this moment on, the character accompanies the visitor throughout the visit, jumping between the mobile PDA and fixed kiosk displays as necessary. The choice of a character implicitly entails the choice of a presentation perspective. In our prototype a lady (Fig. 2) presents historical details about life and activities in the Middle Ages, while a painter is more focused on artistic details, pictorial techniques and restorations. At any time, visitors can come back to the stationary device to switch characters. When this happens, an animation makes the switch explicit: the character disappears from the PDA and hops over to the stationary device (as shown in Sect. 3.1)
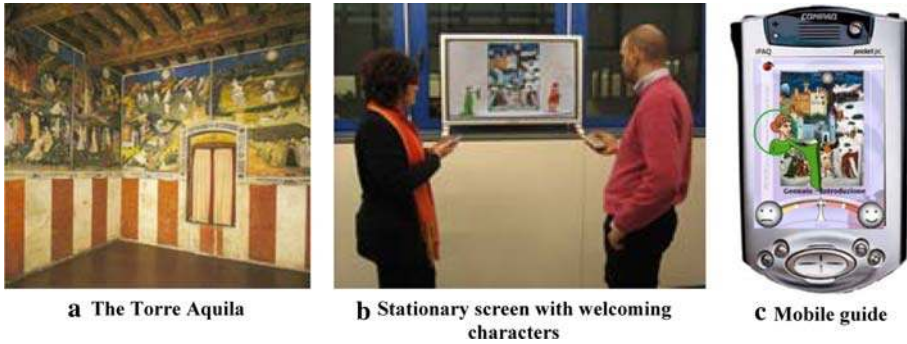
a The Torre Aquila   **b** Stationary screen with welcoming characters   **c** Mobile guide

Fig. 1 PEACH interaction scenario

Fig. 2 The life-like character (in the form of an aristocratic woman) first presents a static image and then announces and starts the presentation of a video clip

During the visit the system shows presentations about details of frescoes that the visitor seems particularly interested in. These presentations include multimedia documentaries, built dynamically and adaptively according to the visitor's interests, as derived from her behavior or explicitly communicated. Observation of behavior includes stops at specific exhibits, time spent in front of an exhibit, presentations the visitor has been exposed to and so on. Explicit feedback to the PDA is expressed through the *like-o-meter*, an interface component that supports the visitor in expressing her degree of "liking" toward the current presentation.

Figure 3 shows a representative snapshot of a dynamically produced video clip that describes an overview of what a visitor sees while looking at presentation for the January fresco. The first still (a) provides a general view of the entire fresco and describes its principal subregions. The video then zooms into one of those regions, a snowball fight

Fig. 3 Sequence of stills from an example video clip

between nobles (b) as it describes via audio what the visitor can see in that area of the fresco. Next, the video pans upward and to the left while zooming slightly out to show the next area, a castle with a rose garden (c). Then, in accompaniment with the audio, the video zooms in to show details about the windows of one of the castle buildings while explaining that they are so highly detailed that they allow curators to precisely identify which castle in the region it depicts (d). Finally, the video zooms back out to show the full castle (e) and then pans to show the third area of the fresco where hunters and their dogs are trying to track foxes through the snow.

A camera attached to the PDA is available for the user to take shots of interesting details. When used this has the effect of transmitting to the system information about her focus of attention (via custom image recognition software) causing the system to provide additional specific presentations on the focused details.

In general the information provided by PEACH is tailored to the user interest and the context of the visit. For example, in the course of the visit the system refers to frescoes that have already been seen and compares them to the current object. As Fig. 4 shows, each fresco contains an enormous amount of detail (each is nearly 3 m tall), and because they are thematically linked by months, seasons, people, architecture, etc., they lend themselves well to cross-reference and comparison. For example, one typical tool to personalize the information is a

Fig. 4 Frescoes representing the months of January, July, and November

dynamic reference to previously visited frescoes. The system thus uses phrases like: "As you have already seen in the January fresco" to enhance the cohesion of the information presented.

At the end of the visit, when the visitor stops to return the PDA, a report generation system directs a stationary printer at the exit to produce a personalized report of the visit just completed, choosing details the visitor was interested in, arranging those details into overarching themes, and offering specific suggestions for further insight (for instance via the internet) and for future visits to both the Torre Aquila as well as other nearby museums.

2.1 Two scenarios of visitors interacting with PEACH

To better illustrate the adaptivity of the system and provide a basis for specific discussion on user modeling, we present two scenarios of different hypothetical visitors who nonetheless are typical of visitors we observed in the Torre Aquila. In the two scenarios we purposely have visitors follow the same Months in the same order, to emphasize the system's flexibility even when their physical paths are the same. But of course visitors are totally free to move as they please and their presentations and reports will vary accordingly.

Once a specific character is chosen, visitors begin to see a series of short, interconnected multimedia presentations depending on the exhibit that the visitor is standing in front of, and, as the visit progresses, her previous explicit feedback to the system and behavior. The user model is updated whenever the visitor interacts with the system or changes position, and becomes the basis of future presentations and of the personalized final report.

Scenario 1: Visitor 1 selects the character representing artistic perspectives (the painter). At this point the system has already pre-selected all the information tagged with such a perspective, which includes such details as the physical state of the fresco, comparisons to other renaissance frescoes and the particular styles of clothing worn by peasants and noblemen. Visitor 1 first arrives at the fresco representing the month of January, and obtains a video presentation that describes its structure and the meaning of its individual elements: a castle overlooking a rose garden, nobles having a snowball fight, and two hunters with their dogs tracking foxes through the snow. The visitor then sees the second presentation shown in Table 1. For the sake of clarity, the video clips are presented here with stills of key frames, each coupled with the transcription of a portion of its accompanying audio and the description of associated camera movement.

Table 1 Snowball fight: standard presentation

	Below the castle a group is playing a snowball fight. This is the first snowball fight seen in Western painting … *[Slow zoom from the whole fresco to the detail of the snowball fight at the bottom]*

Table 2 Snowball fight: additional information from an artistic perspective

	The painting of the six figures was worn, and so much of it was repainted, except for the clothing of the two ladies on the left where you can still see some original paint. *[Pan and Zoom to the 3 characters on the left of the snowball fight scence]*

Table 3 Artistic presentation of the hunting scene

	The figures of hunters, dogs, and the foxes are all original and well conserved much more so than the trees, which in part have lost their original paint. *[Fast fade-out to the whole picture and slow zoom toward the hunting scene in the top right corner]*

The visitor likes the presentation about the snowball fight and provides positive feedback through the *like-o-meter* indicating an interest in hearing further information. The system then dynamically composes the presentation in Table 2.

This choice of the system is influenced both by the initial choice of the character (the painter) and by the visitor's interaction with the PDA.

After the presentation on the hunting scene, the visitor registers interest in the topic, and thus the system selects the presentation in Table 3.

Since each presentation is tied to a domain knowledge base which specifies for instance that the snowball scene consists of nobles engaged in leisure activities while the hunters here are peasants at work, expressing interest in the hunting scene not only causes the user model to record interest in topics such as winter and hunting, but also related topics from the KB such as peasants and social class. The visitor now proceeds to the February fresco, which depicts a tournament where ladies of the court are watching jousting knights on horseback

Table 4 Artistic presentation of the tournament scene in February

	The players of the tournament have been repainted less than the ladies watching from the castle wall above. Many of the original colors have disappeared, especially those of shields and armor. *[Zoom to the leftmost knights in the tournament scene]*

Table 5 Artistic presentation of the ladies in February, with explicit reference to previously seen material

	Unlike the ladies in January that you have seen before, these have been substantially repainted. The original paint has totally disappeared. *[Slow fade-out of the tournament scene and slow fade-in of the ladies (the cross-fade here is used to signal a change of topic)]*

while in the space below a blacksmith in his workshop is repairing the weapons used by the knights. When the visitor halts a presentation on the blacksmith, the knowledge base again connects this interaction with, among other topics, the theme of social roles and work associated with it. Thus while the user model records less interest in the occupation of blacksmith and weapons, it also records lower interest in the peasant class, merging it with that same bit of knowledge stored during the January video.

Visitor 1 does, however, show interest in the tournament scene in the center of the February fresco. The system now elaborates the presentation in Table 4.

During the presentation of the ladies watching the tournament, the visitor expresses interest. The system selects the presentation shown in Table 5.

In this example, effects of the adaptation mechanism are clearly noticeable. Adaptation is possible because the system has kept track of the previous presentations (among them one about aristocratic ladies), and the visitor has chosen the artistic perspective and has shown a positive interest in the scene of ladies in February.

Passing next to the April fresco, the visitor watches all of the detailed presentations about springtime agricultural activities like the plowing and sowing depicted there, but without special interest.

However, visitor 1 also skips the scenes of noblemen and women tending flowers or walking about in new spring clothes. Next, the visitor continues to the July fresco and requests further presentations about summertime farming activities like mowing hay and fishing while ignoring its hunting scene involving nobles with hawks carried by peasants on wooden frames. Finally, Visitor 1 moves to the November fresco and requests presentations about all the major regions of this fresco. However, she stops further presentations about the peasants driving pigs to the town market but likes the focus on the scene where a bear and her two cubs are being hunted, and gives positive feedback to indicate interest. She continues to watch another two

Table 6 Two paragraphs of the report for Visitor 1

...

Winter activities are very common both for nobles and peasants, as you saw in the frescoes of January and February. These are mainly leisure activities for nobles like the January snowball fight and the tournament scene where noblemen are jousting and noblewomen are watching from above. Interestingly, all three of these scenes have had their original paint damaged and then touched up at some later date.

Hunts were popular with noblemen in all seasons, as you can see from the hunt with falcons in July and for bears in the mountains in November. Bears coexisted very uneasily with humans in the middle ages as people began to cultivate more land and raise domesticated animals, like the pigs being driven to market at the bottom half of November. November also contains additional damage to the paint of the fresco both for the pigs as well as the intermediate scene of men and horses encamped on their way to the hunt.
...

Table 7 Historical perspective about the snowball fight

Engaging in snowball fights was a very common leisure activity of noblemen and women during the winter. They used to play this game during festivities.
[Slow zoom from the whole fresco to the detail of the snowball fight at the bottom]

presentations describing the bear and its alpine environment as well as interactions between humans and wild animals.

At this point, the visitor ends her visit and receives a report of her visit, an excerpt of which is shown in Table 6.

Scenario 2: Visitor 2 "adopts" the historical character - the lady - instead. Visitor 2 then initially behaves like Visitor 1: she stops in front of January, watches the general presentation about the fresco and the description of the snowball fight (see Table 1). She gives positive feedback through the *like-o-meter* and the system dynamically composes the presentation reported in Table 7, which describes the same scene but geared more towards the historical perspective.

As with Visitor 1, she gives positive feedback during the presentation of the hunting scene. Since Visitor 2 has chosen the historical perspective she is presented the video in Table 8, whose make-up depends both on the character (the Lady) and the positive feedback just given.

Visitor 2 next also moves to February. She listens to the general presentation of the fresco and then shows special interest in the tournament scene by giving positive feedback, which Visitor 1 did not do. The system now composes the video in Table 9 which is dominated by the historical perspective, but the interest is registered in the user model and exploited later.

Visitor 2 now continues on to the April and July frescoes, behaving similarly to Visitor 1 until she arrives at November. Here, she seems less interested in hunting as she stops

Table 8   Historical information about the hunting scene

	For members of medieval society hunting was very important. Although it is a leisure activity, it was also very important for the economy. Moreover, it was a way to defend cultivated land and cattle from wild animals. *[Slow pan toward the hunting scene in the center of the fresco; here no cross-fade is used, so as to signal topic continuation]*

Table 9   Historical presentation of the tournament scene

	Tournaments have faraway origins, related to the conflict between French and Arab peoples. The adventures of the protagonists were usually handed down orally. *[Slow cross-fade to the tournament scene; a cross-fade is used here to signal that background information is going to be provided]*

Table 10   Two paragraphs of the report for Visitor 2

...

 Winter activities are very common both for nobles and peasants, as you saw in the frescoes of January and February. These are mainly leisure activities for nobles like the January snowball fight and the tournament scene where noblemen are jousting and noblewomen are watching from above. Tournaments were very popular, not only because they allowed people to enjoy themselves outdoors in the late winter months, but they also prepared knights for battle in spring and summer.

A city's walls and gates were important places in the medieval era, where events and festivals like tournaments (seen in the February fresco) were held due to lack of space in the city, and where strangers and farmers living in the countryside had to pass through to trade or continue on their journeys. There were many economic activities based outside of the cities, ranging from farming to hunting, to raising domesticated animals like the pigs that you can see being driven to market here at the bottom of the November fresco. The detailed scenes of farmers plowing, sowing, raking and harvesting are some of the most important elements of the April and July frescoes since they give us clues into the daily lives of the majority of the population in medieval times.

…

the presentation by giving extreme negative feedback about the bear hunt where Visitor 1 expressed interest and in turn gives positive feedback during the presentation about the pig herder driving his swine through the city gates which Visitor 1 disliked.

At the end Visitor 2 receives a visit report which is different from the previous report for Visitor 1 due both to the character and perspective initially chosen (artistic vs. historical) as well as individual behavior and choices recorded during the visit itself. Note that the report generator, searching for commonalities in the knowledge base between the tournament scene in February and the November herding scene, finds such a common element and makes it the topic of the second paragraph in Table 10.

In the following Sect. 3 and 4 we illustrate the two adaptive presentation applications that take part in the scenario described above. Section 5 illustrates the user modeling and adaptation component of the overall system. Section 6 summarizes the underlying software architecture and Sect. 7 discusses evaluation.

3 Distributed multimedia presentations

A museum visit is a personal experience encompassing both cognitive aspects (e.g., the elaboration of background and new knowledge), as well as specifically emotional aspects (e.g., satisfaction of interests or fascination for the exhibits). Various media and resources can be used to produce presentations that engage the visitor and help her appreciate the real thing.

One of the key issues we have worked on is the seamless interleaving of interactivity with mobile and stationary devices. The main requirement is a guarantee of coherence throughout the visit. On the mobile device, personal video clips are dynamically adapted to context, while on larger stationary screens distributed throughout the museum – which we call *Virtual Windows* – further background material and additional information is provided. A virtual presenter follows the visitors on their tour and gives advice on both types of devices as well as on the museum itself.

3.1 The role of life-like characters during presentations

User evaluations (van Mulken et al. 1998) have shown that the introduction of life-like characters makes presentations more enjoyable and attractive (something that we regard as especially important to keep younger visitors engaged). As seen in Sect. 2, in PEACH we have introduced life-like characters that play the role of accompanying agents, ready to move over to the mobile device or to jump on the Virtual Windows, in order to provide continuous assistance and continuity to the presentation. The characters help in solving problems like how to reach a certain exhibit, and producing explanations. The use of life-like characters on portable devices has to be weighed carefully due to the small dimension of such displays.

Nevertheless, there are specific roles that a properly designed character can play on a mobile device to improve the level of engagement with the presentation. In particular, following the TV- metaphor, two main roles can be recognized: the *presenter* and the *anchorman*. When playing the role of the *presenter*, the character introduces new media assets and uses pointing gestures. When playing the role of the *anchorman*, the character just introduces complex presentations without interfering with them any further. Although simpler than the presenter, the role of an anchorman can help the visitor understand many different presentations, providing a context in which they all make sense. In its role as an anchorman the character also supports the seamless integration of the mobile devices' small screen and the large screen of the Virtual Window. Like a TV-presenter who walks around the studio to present different content, the character is able to move between the mobile device and the Virtual Window. Besides the specific role that the character may play, it is also a metaphor for the actual interests of the visitor. By providing different characters and giving the visitor the choice between them, the different views on the exhibits are transparently conveyed and selected.

For example in the scenario described in Sect. 2.1, we introduced an aristocratic woman (see Fig. 2) for a historical and social point of view and an artist for an artistic and technical orientation (Fig. 5).

On the mobile device we have experimented with different embodiments for the virtual character. While a full-sized character provides the most flexibility in terms of gestures and movement, it also unfortunately covers large areas of the screen, something that is not suitable
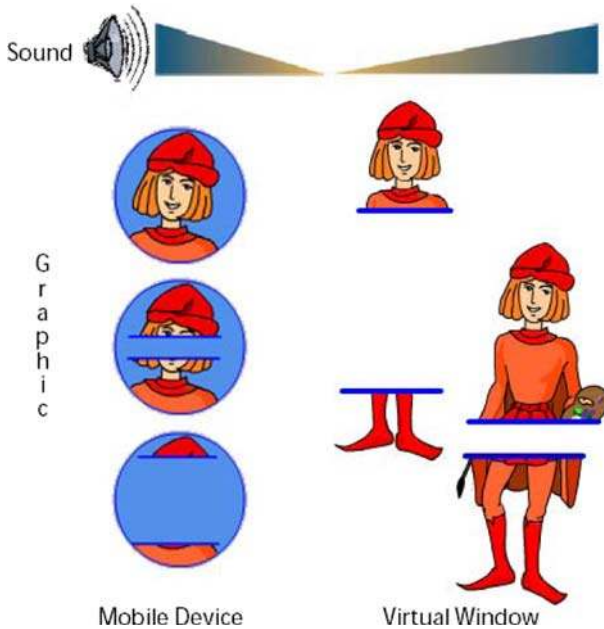
Fig. 5 Key frames of the transition between the mobile device and the Virtual Window (the beam-effect)

for the small displays of mobile devices. For this reason, we have decided to use an iconic representation instead, displaying only the character's head (see Fig. 5). This choice restricts us to a limited set of character animations and gestures, i.e. small head and lip movements and simple pointing gestures, like the one shown in Fig. 2.

3.2 Presentations on Virtual Windows and transitions between devices

The Virtual Window provides visitors with in-depth information on interesting topics. It has enough resolution to allow the full use of graphics, animations and video clips of all kinds. When visitors approach a Virtual Window, their personal presentation agent will transit to the Virtual Window, where it appears full-sized. In order to detect the visitor's relative distance to the Virtual Windows, each of the windows is equipped with two infrared beacons of different ranges. When visitors approach a Virtual Window for the first time, the presentation agent, in its role as anchorman, proactively informs them about the Window and how to make use of it.

If the visitors are close enough, the presentation agent starts to disappear from the mobile device and to reappear on the Virtual Window. The transition from one device to another is emphasized with sounds and an animation. Key frames of such an animation are shown in Fig. 5. This *beam*-effect is used to guide the visitor's attention towards the Virtual Window, where they find the personal presentation agent continuing the presentation. Once the presentation agent is on the Virtual Window, the visitors can continue to coherently interact both with the agent and the presentation. The presentation agent is playing a more active role while guiding the visitor through the presentation on a Virtual Window. Sophisticated gestures and animations thus lead to a much more believable appearance. Before leaving a Virtual Window, the visitor may also change her accompanying character, who after another transition, automatically reappears on the mobile device.

We have also looked at the situation where more than one visitor approaches a Virtual Window at a time. This poses special requirements on interaction and the presentation of information, because several users have to share screen real estate, presentation time and input channels of the Virtual Window. The first problem that has to be solved is how to identify the interaction of a particular user with the Virtual Window. Traditional touch screen solutions fail, since they are not able to map interaction acts to particular users. We took an approach that takes advantage of the fact that each user is in possession of a PDA. When visitors approach a Virtual Window and the character has migrated, as explained above, the PDA can be used to remotely control functions on the Virtual Window. Visitors simply use the touch screen of the PDA to select, start, skip, or stop presentations (Kruppa and Krüger 2003). Of course such an interaction scheme does not work very well if several visitors try to interact simultaneously with the presentation on the Virtual Window. For these situations we have developed a specialized voting mechanism, which collects all topics that are of interest for all present visitors from their user models. Visitors can then review the topic list on their PDA and cast a vote for one of the topics. Afterwards, the highest ranked presentation is selected and presented on the Virtual Window. We have experienced that weighting the votes makes sense to prevent a draw when only a small number of votes has been cast. Assigning the first visitor that arrives the highest vote reflects the fact that this visitor is in "possession" of the Virtual Window.

Another way to solve the problem of multiple interests is to generate a common presentation reflecting all interests of present visitors. We have achieved first results in this respect (Krüger et al. 2002). However, producing interesting and engaging presentations for heterogeneous user groups is far from trivial and is the subject of future work.

3.3 TV-like presentations on the mobile device

Video documentaries represent an effective means of presenting information about a work of art because they rely on the synergy between auditory and visual communicative channels to accompany the audience in a multi-sensorial exploration of the various physical details and the historical/artistic value of the work. Consistent with the PEACH *leitmotiv* that the exhibit remains the real protagonist of the visiting experience, camera movements can provide visual support to the audio commentary by anticipating or following details close-up, or can increase audience engagement during the information acquisition process by emphasizing the rhythm of the presentation. Automatically creating a video documentary to be shown on the PDA screen as the visitor moves around the exhibit, and which is personalized for each user according to preferences and interests, as well as current and past interactions with the system, entails solving several problems: selecting salient (visual and conceptual) information from a large repository, sequencing the text and film snippets in a coherent and meaningful way, and realizing the synchronization and the playback of the intended audio and visual effects. In PEACH, we experimented with two possible approaches to generating personalized video documentaries, with different trade-offs in terms of costs and flexibility: (i) a generative approach, that aims at the dynamic generation of the commentary starting from deep linguistic representation, where the generated text is then synthesized and synchronized with a dynamic animation of still images to get video sequences; and (ii) a simpler approach assembles pre-prepared chunks of audio/visual material.

In both approaches, principles derived from cinematography, with heuristics and constraints on how cameras should move (pan, zoom, tilt) or which transitions should be included to improve efficacy (fade, cut), have been considered as the essential backbone for defining planning strategies.

3.3.1 Automatic generation of video clips

The most ambitious of the two approaches we have followed consists of automatically generating on the fly the text of the commentary using deep linguistic representations, synthesizing it, and synchronizing the resulting audio with a video obtained by suitably assembling an animation of 2D still images of the described exhibit.

The benefits of this approach are multi-faceted: (1) deep Natural Language Generation (NLG) allows for a finer-grained control in the generation of referring expressions and other linguistic devices that help improve the naturalness and cohesion of the text (e.g., aggregation), and facilitates the generation of personalized texts which are adapted to the current interaction context and user preferences; (2) as in many other application domains, users in the museum scenario come from many countries and speak their own native languages, making multilingual generation (a laborious task for templates) very important; (3) the underlying architecture can be shared with other application tasks over the same domain (e.g., direction giving or report generation for museum visits, as described in Sect. 4) or other projects requiring generation, reducing the intensive costs of creating deep linguistic resources and domain models; and (4), in a more general perspective, deep NLG also allows for the generation of tags, e.g., marking anaphoric references or the rhetorical structure, that direct the speech synthesizer to produce higher quality output, helping predict accent size and contours as well as accent placement (though this was outside the scope of the PEACH research areas).

Within PEACH, we implemented an engine that takes as input a series of still images and the knowledge base containing information about those images, as well as information about their domain in general (Callaway et al. 2005a). The component selects and organizes the content to be conveyed and produces textual descriptions using standard deep NLG techniques, while a video planner submodule, taking into consideration the discourse structure of the commentary, plans film segmentations. This means defining shots (i.e. video units realized via a continuous movement of the camera over different regions of the same image), selecting camera movements and transition effects between shots. The output of the engine is a complete script of a "video presentation", with instructions for synchronizing images and camera movements accompanied by the audio commentary synthesized from the generated text. One of the chief novelties of this work is the use of rhetorical relations (Mann and Thompson 1987) to help provide structure to both the image sequences and the spoken part of the script. As depicted in Fig. 6, the architecture of the system is based on a bipolar cascade model. The two processes, the generation of the verbal content and the planning of the video script, can be partly performed in parallel. Parallelism allows for improvements in speed (which can be very important if the generation of personalized videos is embedded in an interactive system, Rocchi et al. 2004) and in modularity, allowing both reusability of components and the possibility of experimenting with different approaches to cinematography or text generation.

The NLG cascade is organized following the standard architecture proposed in (Reiter 1994). In the first phase, relevant content of the document is retrieved from the knowledge base and organized into a discourse plan structured as a rhetorical tree. The User Model is accessed to take into account what the user has already seen and heard, his inferred interests, and language preferences. The content selection and organization approach adopted in our text planner is based on the assumption that in descriptive texts/commentaries, the conventions on how information is typically presented play a major role. For example, when describing complex depictions on a painter's canvas, the description (as well as the corresponding image sequence/animation) needs to reflect the spatial organization of details and their salience in motivating the painter's choices. These "patterns of appropriate ordering" are widely known
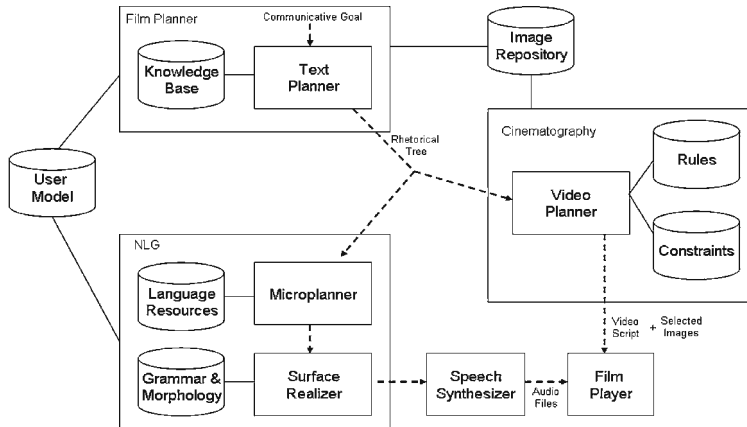
Fig. 6 System architecture for dynamic documentary generation

in the NLG community as *schemas* (McKeown 1985) or the *Generic Structure Potential* (GSP) of a text (Halliday and Hasan 1985). Schemas can be defined at different levels of abstraction and can contain calls to other finer-grained schemas. Each schema comes with an applicability condition that specifies the discourse context, the user modeling conditions and the required information in the knowledge base that allow its use. We implemented a schema-based text planner that chooses the best schema satisfying the current communicative goal and performs a decompositional expansion of the schema calls in its body, ultimately producing a rhetorically annotated discourse tree whose leaves are predicates extracted from the KB to be linguistically realized.

As a clarifying example, let us consider scenario 1 described earlier in Sect. 2.1, and in particular how the text shown in Table 5 is composed. The object in focus in the previous presentation is represented by the ladies watching the tournament in the February fresco. From the user behavior the system decides to prepare an additional presentation. As recorded by the User Model component, at this stage of interaction a high level of interest has been inferred on the scene of the ladies in February and on the concepts related to the artistic perspective (colors, painting techniques and preservation conditions). To enforce smooth discourse continuation, the text planner looks for schemas that select from the knowledge base information about those ladies not previously mentioned. Preference is given to schemas addressing concepts and instances with high levels of interest. In our example, a schema extracting predicates about the painting conditions of the ladies from the knowledge base is well-suited to the current narration context ("The ladies are substantially repainted. The original paint disappeared totally"). However, a more specific schema is available, that also includes a comparison with the painting conditions of a similar detail already seen in the January fresco and this is preferred, so as to improve the overall coherence of the visit ("Unlike the ladies in January that you have seen before, these are substantially repainted. The original paint disappeared totally.").

After the text planning stage, the microplanner applies pragmatic strategies to decide how to realize nouns (i.e., whether to use anaphora or full descriptive noun phrases, etc.) and verbs (i.e., decide the tense, the aspect and so on). Finally, a surface realizer completes the process by selecting closed-class words and enforcing syntactic and morphological constraints according to the grammar of the target language. In our system, the cascade adds English and Italian speech synthesizers as well in order to produce verbal commentary for each video.
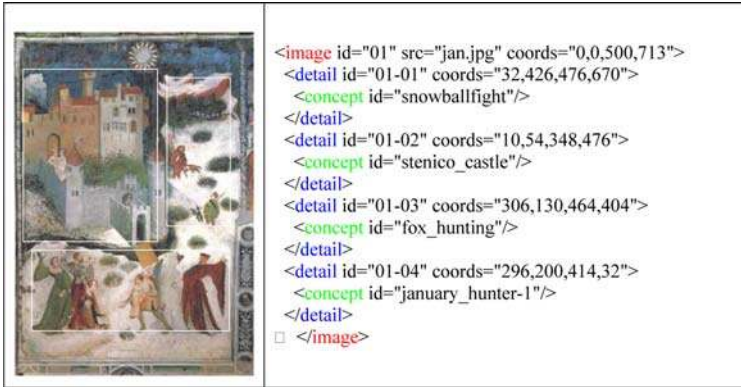
```
<image id="01" src="jan.jpg" coords="0,0,500,713">
  <detail id="01-01" coords="32,426,476,670">
    <concept id="snowballfight"/>
  </detail>
  <detail id="01-02" coords="10,54,348,476">
    <concept id="stenico_castle"/>
  </detail>
  <detail id="01-03" coords="306,130,464,404">
    <concept id="fox_hunting"/>
  </detail>
  <detail id="01-04" coords="296,200,414,32">
    <concept id="january_hunter-1"/>
  </detail>
□ </image>
```

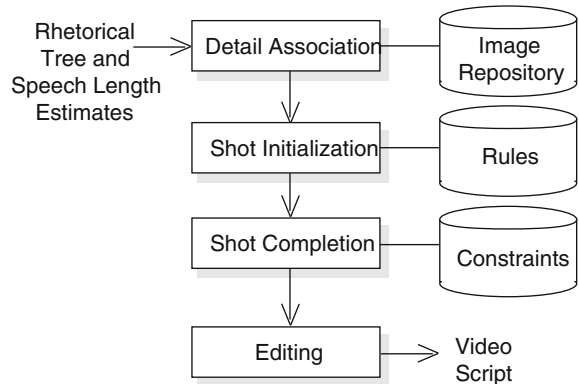**Fig. 7** An image and its annotation

The Video Planner implements a mechanism that automatically computes a full script for a video documentary. It starts from the discourse plan annotated according to the RST annotation scheme (Mann and Thompson 1987) generated by the Text Planner, and the estimated duration of the synthesized speech. In addition, the video planner also takes as input a repository of annotated images containing information about general features of each image (height, width and source file location) as well as information about the details it contains, and relevant subregions of the image illustrating particular topics (see an example in Fig. 7).

The system relies on a core of rules and constraints. They encode the rhetorical strategies that are the basic resource for selecting appropriate images, designing the presentation structure, completing each shot, synchronizing the visual part with the audio commentary while avoiding the "seasickness" effect (back and forth motion). The rules, fired by a forward chaining mechanism, are context sensitive. They are also dependent on: (i) the rhetorical relations among the text spans; (ii) the geometric properties of images selected from the information repository, and (iii) the matching of topics among segments and images. This ensures that the personalization of the audio commentary driven by the user model is reflected also in the visual part of the presentation. Figure 8 shows the structure of the engine.

The video planning consists of four phases:

- *Detail Association*. A detail is associated with each segment of the commentary. In this phase, the system assigns one or more exhibit details to each segment of the commentary. This operation is performed by searching the image repository for details with the same topic as the segment. The preferred heuristic is to select the detail with exactly the same topic(s) of the segment. If this is not possible, search rules look for details which subsume the topics mentioned in the commentary.
- *Shot initialization and structure planning*. The details identified in the previous phase are grouped for potential inclusion in the same shot and a candidate structure for the final presentation is elaborated, starting from the rhetorical structure of the commentary. (The result of this phase can be changed because its processing is iterative). The processing is guided by a set of rules that are fired when particular configurations of rhetorical relations are matched. For example, an elaboration or sequence relation signals a smooth transition from the current topic to new information that is strictly related to it. It is thus preferable to aggregate segments in the same shot and to exploit camera movements to accompany the topic shift. There are cases in which the structure planned in this phase is revised during successive stages of computation. For example, to avoid the "seasickness" effect,

Fig. 8 Architecture of Video
Planner component

Rhetorical Tree and Speech Length Estimates → Detail Association — Image Repository

Detail Association ↓

Shot Initialization — Rules

Shot Initialization ↓

Shot Completion — Constraints

Shot Completion ↓

Editing ⇒ Video Script

the system can apply constraints and then modify the previously planned structure by adding new shots.

- *Shot Completion*. Camera movements between details in the same shots are planned. Constraints are considered in order to avoid "inconsistencies". In this phase, the engine incrementally completes each shot by expanding and illustrating each of its segments. When a candidate move is proposed, the system verifies that it is suitable according to the list of previous camera movements and the constraints imposed over that category of movement. These constraints encode the cinematographer's expertise in selecting and applying camera movements in order to obtain "well-formed" shots. For instance, when a zoom-in movement is proposed where the previous movement is a zoom-out, the system discards the planned sequence as unacceptable.

- *Editing* Transition effects among shots are selected according to the rhetorical structure of the commentary. This is the phase in which the engine chooses the "punctuation" of the presentation. In order to reflect the rhythm of the discourse, the choice of transition effects is guided by the rhetorical structure of the commentary. The system retrieves the last segment of the shot currently considered and the first segment of the shot to be presented next and plans the transition. A sample rule is "If two segments are linked by a relation of type elaboration then a short cross fade applies". The rules for editing have been defined by combining our insights about the discourse structure with the knowledge extracted from our interviews with experts. A final tuning of the rules takes into account conventions of the typical employment of transition effects in the field of cinematography (Arijon 1976). For example, fade effects are appropriate for smooth transitions when the focus of interest changes, but the new topic is related to the old one, as in the case of elaboration or background.

The output of the video planning process is a complete script for the video and audio channels encoded in an XML-based markup language.

Our work provides a new contribution to a tradition of research in automatic production of dynamic images. One of the first case studies of generation of "motion presentations" is the work of (Karp and Feiner 1993) Among others, animated presentations have also successfully been employed in the illustration of technical devices (Butz 1997); camera-planning techniques were proposed in (Bares et al. 1998) and (Halper and Oliver 2000); (Friedman and Feldman 2004) describe a non monotonic reasoning approach to produce 3D animated movies from screenplays.

```
<movie title="Giugno - La cittadina" exhibit="june">
<shots>
  <shot id="giu-cit-abs1" file="${resources}giu-cit-abs1.swf" />
  <shot id="giu-cit-abs2" file="${resources}giu-cit-abs2.swf" />
  <shot id="giu-cit-abs3" file="${resources}giu-cit-abs3.swf" />
  <shot id="giu-mag-citta-compS" file="${resources}giu-mag-citta-compS.swf" />
  <shot id="giu-mag-citta-compV" file="${resources}giu-mag-citta-compV.swf" />
</shots>
<editing>
  <display shot="giu-cit-abs1" />
  <rule>
    <UM-expression>!$user.shotSeen("mag-natura-abs2")</UM-expression>
    <display shot="giu-mag-citta-compS" />
    <display shot="giu-mag-citta-compV" />
  </rule>
  <display shot="giu-cit-abs2" />
  <crossfade shot="giu-cit-abs3" />
</editing>
</movie>
```

Fig. 9 Sample XASCRIPT video template

3.3.2 Adapting video clips to the situative context

The extreme flexibility in content and textual personalization, as well as in visual rendering, implemented by the approach to video generation based on NLG techniques and camera animation is counterbalanced by the intensive costs of creating deep linguistic resources and domain models. When a museum guide system needs to be ported very rapidly and at reduced costs to different domains (as is in the case of temporary exhibitions), less flexible, but still effective solutions, can be developed based on the substantial reuse of pre-existing texts (and related images or audio commentaries) prepared by domain experts.

In PEACH this was realized by offering to a human author a language for composing adaptive videos from existing multimedia material. A flexible mark-up formalism was defined to write video templates, i.e. instructions on how to compose video shots with existing audio commentaries, including choice-points to be resolved at assembly time according to visitor preferences and other user-dependent parameters (Rocchi and Zancanaro 2004). Once a clip is requested for a given exhibit, the Adaptive Video Engine chooses a suitable template and elaborates it according to the current user model to produce an instance of a video clip that is personalized for the given visitor in the given context. There are two ways to affect the presentation with user-dependent features: varying (i) the actual selection of shots and (ii) the choice of transition effects between shots.

Video templates are written using an XML-based language called XASCRIPT (see Fig. 9 for a sample template written in XASCRIPT).

In creating a template, the author first defines a shot repository (specified within `<shots>..</shots>` tags), i.e., a list of existing shots that can be used to assemble the actual video. In the second part of the template, the editing is specified, i.e., how pieces of information (shots) are presented to the visitor. The choice of transitions affects the flow of discourse and highlights the passing from one piece of information to another. In classic text-based hypermedia, transitions from one document to another might not be important for user understanding. In the case of video-based hypermedia, they are crucial, for they underline the rhythm of the presentation and signal the (semantic) "closeness" of the scenes to which they apply. For example, a short cross-fade typically signals a smooth transition of the narration to a related topic.

XASCRIPT adaptation rules can be inserted in the editing section to express alternative ways in which shots should be sequenced, which transitions apply or how long transitions should last. An adaptation rule is a <condition, action> pair, where conditions include requirements for the correct firing of the rule and actions are pieces of documentary
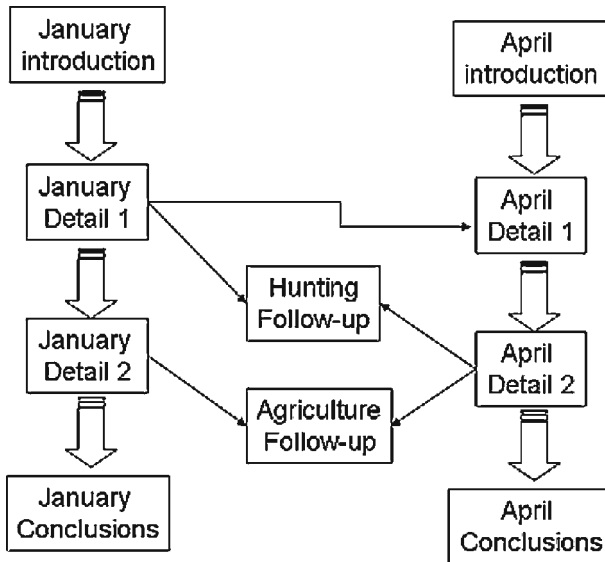
Fig. 10 A simplified network of video clip templates

(transitions) or other rules. XASCRIPT supports two types of resources for defining the conditions: (i) user-model features (UM-expression) and (ii) editing features (EDIT-expression). An UM-expression amounts to a check over the set of features encoded in the user model (described in Sect. 5). For example, in the template in Fig. 9, a UM-expression checks whether the visitor has already seen the shot named "mag-natura-abs2". EDIT-expressions are instead conditions over the editing of the current movie that include dependencies among content units (e.g., when the selection of a given shot requires that another shot has already been included in the current presentation), or constraints between presentation forms (e.g., "if the last transition is not a cut, the next shot can fade-in"). The combination of these two conditional resources allows us to define fine-grained templates, providing a flexible mechanism that supports both content adaptation and dynamic selection of transition effects.

The human-authored templates are organized as a network of semantically related content (see Fig. 10), which is exploited by the UM component to propagate the interests of the visitor, hypothesized according to her behavior and interaction during the visit (see Sect. 5.3).

Figure 10 shows a simplified example of a template network. Each node of the network is a template from which different video presentations can be composed. Each template may contain micro rules, for example for choosing from different linguistic expressions, according to the visit history. Each node is connected to others, where connections are the semantic links over which the interest is propagated, as detailed in Sect. 5.3. For example, if the visitor expresses interest in January during the introduction such information is propagated to all the templates tagged with January.

Templates are of different types: (i) introduction (when they contain overview information); (ii) presentation of details; (iii) conclusions (these are played when the system has exhausted its information about a particular exhibit; the main purpose of these templates is to suggest that the visitor move on to another exhibit) and (iv) follow-ups (which provide insight on general topics). Follow-ups are not strictly related to a particular exhibit; rather they provide connections between different exhibits.

Fig. 11 Three designs of the mobile multimedia guide user interface

The Adaptive Video Engine component is able to customize video clips by selecting and instantiating the most appropriate template. This is accomplished by navigating through the template network (considering what the user has already heard and her assumed interests), choosing the right template and by instantiating its content using adaptation rules.

3.4 Design of user control and delegation of the adaptation process

Explicit user feedback to a presentation is an important component in our adaptive guide. The interface is a critical element, it can also play a role in transmitting some form of feedback to the user, and had to be studied carefully if we want the user to understand what is going on and act correctly. In the literature it is often emphasized that users should be allowed to be in control of the interaction (Shneiderman 1998; Norman 1998) and studies on adaptive systems (see Sect. 7) have raised the issue of controllability of system adaptivity.

With these premises we put a lot of effort during the design phase on the definition of a widget to help the visitor feel in control of the adaptation process. The basic idea was to develop a widget that supports the visitor in expressing his/her degree of liking toward the current presentation, allowing the system to tune the current as well as the future presentations. The resulting widget, the *like-o-meter* (Fig. 11, right side), is conceived as both an input and an output device.

The basic requirements for the interface design was that the guide had to intuitively enable the user to express her liking of the current presentation topic, and it had to be proactive in order to avoid user disorientation: it should be possible to use the guide even without any explicit input beyond physical movement of the visitor herself.

The first versions of the guide implemented just two buttons: "Wow", on the upper right side of the PDA screen, and "Basta!" (Enough!) on the lower left side of the screen (see Fig. 11, left side). A preliminary heuristic evaluation was pursued to remove major usability problems.

The use of the "Wow" button was intended to reflect the affect of the visitor whenever she is positively impressed by a specific presentation. The "Basta!" button was to be used whenever the visitor is not interested in the current topic. As a side effect, this button stops the current presentation. It is worth noting that a presentation can also be stopped by moving away from the current exhibit and approaching another one. The central part of the screen was used to show the video presentation (or to host the virtual character as explained in Sect. 2).

We conducted a pilot evaluation and a qualitative study that showed that the users did not understand the two-button interface. In general, they seemed to perceive the "Wow" button as the key element of the whole interface, often using it to start the interaction (even if this was not necessary) and to request additional presentations. However they complained about the lack of proper feedback after the "Wow" button was pressed.

User studies also provided evidence that the time necessary to load the presentation was a major cause of disorientation. The subject (i.e., visitor) gets disoriented because she doesn't know what to expect, and doesn't understand what (if anything) she should do. Since the "Wow" button is the only visible thing the subject could act upon, she presses it, probably to get information. When all extended personalized presentations have been exhausted, the user, not knowing the state of the system, was still waiting for something to occur. Again, the only available button was the "Wow" and this led to its continual overuse. Finally, the "Basta!" button was only interpreted as a 'stop' while the intended meaning of "I don't like this" was not understood. In fact, given the conceptual model, the system should not allow the user to take any action, but simply to express her degree of liking of the topic.

User evaluations showed that the presentation should not abruptly stop but invite the user to move to another fresco; it should give feedback about the system status, and inform the user about its estimation of her interests; it should skip uninteresting presentations, and focus on more interesting ones, or suggest moving to more interesting frescoes. Finally, if the visitor does not express any feeling (i.e., if she never presses any button), she should receive a fair amount of information about the museum exhibits.

We then designed a new user interface based on the *like-o-meter* concept (Fig. 11, central image, shows the first attempt at designing the widget). The idea was that (a) the visitor should be able to express how she likes the topic being currently presented and (b) that the display widget itself could be used by the system to express its assessment of the visitor interest in the current presentation. The results of a new small qualitative user study in a lab setting were quite encouraging (Goren-Bar et al. 2005), showing a high degree of understanding and satisfaction by the users. The participants were able to communicate their interest to the system by correctly using the *like-o-meter* and recognized that the positions +1 and +2 on the scale caused more information to be provided. They also understood that when the system starts a new presentation, the highlighted position in the *like-o-meter* is determined by the system. One participant, in particular, clearly noticed that this reflects system interpretation of her previous behavior and modeling of her interests. Given the limitations of this small study, we could not reliably conclude that the delegation metaphor was properly understood by the subjects, though this seems highly likely given the available evidence. In particular, we could not reliably conclude that they fully realized that their expression of interest on a given current exhibit also affected the presentations to come.

Some usability issues also emerged, in particular for what concerns the understanding of the meaning of moderate disliking (i.e. position −1): it was somewhat poorer than that of liking. Apparently, the users came to expect that the expression of a moderate disliking should cause the system to provide less information. Indeed, in our current system, the expression of moderate dislike only changes the user model and does not affect the current presentation. Similarly, they did not expect that the neutral position of the like-o-meter could be selected, and expected that a single button press would have moved the slider from −1 to 1.

Evidence collected in this study drove the design of the final version of the interface shown on the right side of Fig. 11, where the user can touch the faces to raise/lower the rating. This final design was the subject of an extensive quantitative evaluation with the system, as reported in Sect. 7.2. We would like to emphasize that resolving gross interface issues via preliminary, qualitative user-centered design is essential for focusing subsequent large

scale evaluations on the fundamental aspects of the interaction with an implemented adaptive system.

4 Report generation

The PEACH report generation component has the goal of allowing visitors to continue interacting with exhibits even after they have left the museum by creating a written report in either English or Italian. Each report begins with a personalized header (name, date and time of visit), contains a basic narration of their visit in one of the three styles described below (including the particular items and relationships they found most interesting), has color images of relevant artwork placed near the text allowing them to recall those artworks when they read the report at a later date, and at the bottom of the page includes links to additional information on the internet and related nearby museums as a natural follow-up for a specific visitor's interests. This extended interaction allows visitors to learn more about topics of interest and even plan future visits where they can get relevant information they did not initially see due to various constraints such as limited visit time. Thus the current museum visit becomes a single episode in an ongoing coherent sequence.

4.1 What makes a good post-visit report?

Ideally, a visit report should take into consideration the following elements (and perhaps more):

- Factual aspects of the visit, such as the subset of exhibits visited, the ordered sequence of exhibits visited and time spent at different locations, the presentations shown to the user, and user actions in the information space, such as interactions with the mobile device.
- The appearance of the report, combining text, images and possibly additional forms of media, in a personalized manner either on paper or in an electronic form.
- Cognitive aspects related to the exhibits, such as interest in themes, happiness, boredom, etc., which may be determined directly or inferred by user behavior such as long pauses.
- Extra subject-centered aspects, such as persons met, discussions held or additional events that occurred in the physical space.
- Attention-grabbing elements, hints for subsequent reading and visiting (in this and other museums), and current events such as a related medieval festival in a nearby town.

The quality of the report is essential: it should be a memory aid for later consultation, something one can share with others, and an entry point for getting deeper into a subject. It should be short, yet readable and concrete. Details often are important, but only when relevant to that specific visitor.

An important aspect in deciding how to construct the report is its organization—it could be *sequential* following the particular sequence of visited exhibits, preserving the visit path and recording events in a chronological order; it can be *thematic*, by depicting the essential themes the user was interested in but abstracting away time; or it can combine both of these. The content of the report should include details related to the visitor's requests for information, interesting connections between seemingly unrelated exhibits, comparisons between the exhibits the visitor found interesting, and descriptions of related exhibits they would probably have been interested in but did not have a chance to see. It also has to allow some inherent variation and adaptation, so that two users visiting together and walking on the same path but making different choices would still get different reports.

The report should make the visitor want to continue to study the topics depicted by the exhibits and make this study easier by suggesting relevant links to follow up at a later date, additional items to visit (either now or on a future visit) and possible additional exhibits in nearby museums. Because museum visitors typically come from all over the globe, it is desirable to provide them with the report in their native language. Additionally, the final report should be producible as printed hardcopy and handed to the visitor as they leave, or if they wish to provide their email address, emailed to them with hyperlinks generated inline within the text to allow follow-up interactions at home, or as a complete multimedia show at a museum kiosk that combines their interests into a dynamically produced film (Callaway et al. 2005b).

4.2 Report generation in PEACH

Within PEACH, we have tried to address all the challenging issues above, where at the end of a visit to the Torre Aquila with the adaptive tour guide, a custom report is printed and given to the visitor. The report includes an overview of the frescoes and themes seen, additional information about specific topics they appeared to have been interested in, non-obvious relationships to other similar topics they didn't have time to see, and follow-up web links to sites of other nearby museums allowing visitors to continue exploring these topics at home.

The PEACH report generator is composed of an NLG system communicating with the user modeling component, both of which are based on a common multilingual domain knowledge base (KB). The NLG system (Fig. 12) consists of a text planner which knows how to organize text above the sentence level, a set of reporting strategies that can query the user model in order to determine the best high-level approach, and a sentence planner that can create surface forms for individual sentences from linguistic representations in the KB while simultaneously adding images and hyperlinks. The knowledge base contains detailed representations of four out of the eleven frescoes (January, February, April and November) in Torre Aquila and enough information on the remaining frescoes to support automatic comparisons and contrasts for purposes of user modeling and adaptivity. As described in Sect. 5, the domain KB (examples may be seen in Fig. 16) contains a domain ontology, facts about each fresco as a whole and the visual elements they contain, generic concepts which are not visually represented in the fresco but may be used to explain them (like "aristocracy" or
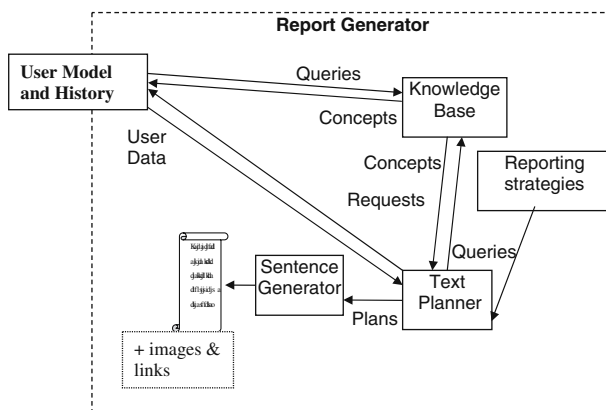

Fig. 12 Report generation component architecture

"winter"), and links between facts to allow for inference (such as comparisons, contrasts and complementary information).

The user model provides inferred information about what level of interest the visitor has in the concepts associated with the presentations she saw. Using an inference mechanism that follows ontological links in the KB, the specific concepts associated with the presentations are augmented by additional, related concepts that extend interest to categories of concepts beyond those that were associated with the presentation seen by the visitor. The accumulated information in the user model is preserved for the visit summary report, generated at the end of the visit.

The Text Planner determines the most relevant information to put into the description and its coherent organization. In the case of report generation, coherence refers to correct selection of topics and details and their placement in sentences to achieve recognizable sequential and thematic structures. Information (content) about the artworks seen is extracted from the KB (e.g., objects like castles or wild animals in a fresco, their positions, and what activities they may be performing) and given either important roles such as the main topic of a paragraph or else that of supporting detail.

For adaptive generation that is highly personalized for a particular museum goer's visit, it is important to ensure a high amount of variation in the resulting text. To achieve such variability, the text planner queries the user model to get the log of the user interactions. For instance, to sequentially describe what the visitor saw, the text planner extracts from the log contained in the User Model the list of visited frescoes and accesses the KB to get a shallow description of the main contents of each artwork to be included in the summary as a reminder of what was seen.

Alternatively, the text planner can retrieve a list of ranked topics from the inferred interest model. The corresponding thematic report might consist of a series of paragraphs describing the top-ranked items in the interest model. To prevent repetition, we also cluster items in the interest model that are semantically related. Which details from the KB to include and how they are ordered are determined by text planning rules which query the user modeler and are implemented at the document planning level with a schema-based approach as described by Callaway et al. (2005b).

For example, as a result of the first scenario in Sect. 2.1, the user model constructs an interest model representing the visitor's heightened interest in winter activities such as the snowball fight and tournament (both activities involving aristocracy). The interest model clusters these local interests to produce abstract topics such as "aristocratic activities" rather than "knights" and "ladies" individually. If well-defined clusters can be created, the text planning component then chooses a thematic report centered around the top clusters in the list. Otherwise, it chooses a sequential or hybrid report (depending on how many clusters can be found) that describes what the visitor saw in sequence, calling up elements from the interest model to serve as either scaffolding topics or extraneous details depending on the level of interest recorded.

The text of the report is created using the language-independent StoryBook deep NLG system (Callaway and Lester 2002). While the text planner takes care of high level communicative goals and varies the report to make sure it is individualized, the generation component focuses on low-level issues, such as adding pronouns, making sure sentences have subjects and verbs, adding morphology to word stems, etc.

Deep NLG has advantages that make it useful in generating extended reports in a museum context: the text can be in multiple languages (in our case Italian and English), produced in high-quality prose, provide for automatic variation at the lexical level, and contain integrated markup (such as HTML, or prosody for TTS). Formatting is particularly useful for report

generation, as it allows for the creation of the report as a webpage, with hyperlinks and images of artworks for each paragraph, allowing the visitor to mentally associate them with the text as they are reading about their experience with it.

5 User modeling and adaptation

PEACH user modeling was required to be "non-intrusive"; hence visitors are not required to fill out questionnaires or to provide any personal information and user modeling is based solely on visitors' observable behavior. As a result, at the beginning of the visit there is no available knowledge about the visitor and there is no detailed user model, only a "lean" user model based on stereotypical information, as will be described shortly. For achieving a high level of adaptivity, the visitor is tracked during the visit by recording her positions (in terms of the visited exhibits), any events that took place, and time spent at every position. These data are recorded by the user modeler component, together with the specific presentations delivered and any feedback (implicit and/or explicit) provided by the visitor. Implicit feedback recorded consists of either normal or abnormal presentation termination (when a visitor walks away before the current presentation ended). Explicit feedback, instead, as described in 3.4, derives from simple explicit interaction with buttons displayed on the interface allowing the user to express her satisfaction with the presentation.

The user model could exploit information already existing in domain knowledge base used for natural language generation and had to support distinct applications (as we have seen in the presentation and report generation systems described in Sect. 3 and 4).

As we will briefly describe in Sect. 6, in PEACH we have invested in a powerful multiagent architecture that can accommodate services in an open ended way. One motivation for this architectural approach is opening new opportunities for user modeling (as well as for integration of additional services in active museums). In principle, there is no longer a need to rely on a single user model: several user models can work (and evolve) in parallel—a stereotypic user model, a knowledge-based, adaptive user model, a template-based approach to interest propagation, a model based on collaborative filtering and more. All of them may play the same "role" of user model and they form an *implicit organization* (see Sect. 6.1) of user modelers for any given individual visitor. Whenever there is a need for information for a specific user, a query can be posed to the implicit organization of user modelers. The query is distributed on a channel where all agents playing the role of user modelers are tuned to. Each and every user modeler can reason about the request and suggest a reply, if available. The user modelers can negotiate among themselves, in order to determine which reply is the most accurate at a given moment and the winning agent provides the reply to the requester as a reply of the implicit organization. The user modelers can compete (bid) for winning the opportunity to reply, or collaborate in order to generate a mutual response, applying the negotiation policy selected by the organization (as presented in 6.1). This may be a better and more flexible solution than relying on a single user model. A simple example can illustrate the benefits of this architecture: at the beginning of a visit, with no information about a visitor; a stereotypic user model may be the only source of user modeling information. Such a model can be invoked following the visitor's selection of a virtual character that can accompany her.

During the visit, information accumulates, and when an assessment of the user state is requested, each and every model can suggest one, together with a certainty factor, based on the confidence the agent has on it. The individual models' assessments may be combined, taking into consideration the individual certainty factors. At a certain point the personal user models may provide better predictions than the stereotypic one and take over completely.

```
<profile>
   <loadUserInterests>
   …
   <interest topic="LEISURE-ACTIVITIES" level="very interested"/>
   <interest topic="HANDICRAFT-ACTIVITIES" level="interested"/>
   <interest topic="ARISTOCRATIC-CHARACTERS" level="interested"/>
   <interest topic="PLEBEIAN-CHARACTERS" level="very interested"/>
   <interest topic="NATURE" level="interested a little"/>
   <interest topic="ANIMALS" level="interested a little"/>s
   <interest topic="UTENSILS-AND-PRODUCTS" level="interested a little"/>
   …
  </loadUserInterests>
</profile>
```

Fig. 13 Stereotypic user model initialization example

This moment of transition can be determined dynamically, based on the performance of the models. Moreover, this situation may be changed dynamically and they may switch back whenever a visitor arrives at a new exhibition/section where for instance the knowledge-based model may have limited or no relevant knowledge, hence the stereotypic may again provide a better prediction.

This dynamic behavior enriches the possibilities for user modeling with very little additional effort, allowing for the application of several different techniques in parallel and to use the best possible composition of services at any given moment.

With the aid of the sample presentations and reports presented in Sect. 2.1 above, we now examine in more detail how the basic user models are initialized and evolve during interaction.

Let us consider in more detail the second visit scenario. The visitor "adopts" the historical character—the lady—as a companion. Then she stops in front of January, watches the general presentation about the fresco and the description of the snowball fight. She provides explicit positive feedback and the system dynamically composes a follow-up presentation.

5.1 Initiating a user model – stereotypic approach

Initial information implicitly provided by the visitor can help in bootstrapping a personal user model or serve as a model throughout the visit. Stereotypes with predefined levels of interest associated with the most representative concepts of the domain model were defined and coupled with the virtual characters the visitor selects to accompany him/her during the visit.

Figure 13 illustrates the information associated with the "Lady" virtual character, representing implicit interest in history and social perspectives of the life in Trento in the Middle Ages. For example, ARISTOCRATIC-CHARACTERS and PLEBEIAN-CHARACTERS are concepts assumed to be respectively of interest and of high interest to the visitor who selects the "Lady" virtual character, hence they are added to the user model with an initial value of "interested".

Having some initial information about the visitor is crucial for good personalization. Besides the solution proposed here, the emerging research in the area of "Ubiquitous User Modelling" is becoming an interesting solution, where a few recent research examples can be found in (Potonniee 2002; Kay et al. 2005a; Niederee et al. 2004; Heckmann et al. 2005). The possibility to use partial user models, already available from previous diverse experiences, in order to bootstrap the ad-hoc user model needed for the museum visit, will greatly enhance the ability of the museum visitors guide system to provide personalized presentation to the visitor and solve the "cold start" problem which is one of the major challenges in such a system (Berkovsky et al. 2006).

```
<userTopInterests userId="p02av" >
…
<interest topic="LEISURE-ACTIVITIES" level="very interested" />
<interest topic="SNOWBALL-FIGHT-SCENE001" level="very interested" />
…
<interest topic="PLEBEIAN-CHARACTERS" level="interested" />
…
<interest topic="ARISTOCRATIC-CHARACTERS" level="very interested" />
…
<interest topic="HANDICRAFT-ACTIVITIES" level="interested"/>
<interest topic="ARISTOCRATIC-CHARACTERS" level="interested"/>
<interest topic="NATURE" level="interested a little"/>
<interest topic="ANIMALS" level="interested a little"/>
<interest topic="UTENSILS-AND-PRODUCTS" level="interested a little"/>
<interest topic="CASTLE-SCENE001" level="interested a little" />
<interest topic="JANUARY-FRESCO001" level="interested a little" />
<interest topic="SOCIAL-CLASS" level="interested a little" />
<interest topic="HUNTER-SCENE001" level="interested a little" />
<interest topic="PICNIC-SCENE005" level="interested a little" />
<interest topic="PLEBEIAN-VILLAGES-AND-FARMS" level="interested a little" />
…
</userTopInterests>
```

Fig. 14 Interests level representation for domain concepts and instances. Bold words illustrate terms used for inference in the example

5.2 Knowledge-based user modeling

The Knowledge-Based User Model draws some ideas from prior work in projects such as AlFresco (Stock et al. 1993), HyperAudio (Sarini and Strapparava 1998) and HIPS (Benelli et al. 1999), regarding user interest representation and interest level inference.

The User Model maintains a list of relevant domain concepts and instances mentioned in the presentations delivered to the visitor and for each of them it associates an inferred level of interest, which is updated according to the user's behavior during interaction.

In contrast to the approach in the work mentioned before, only visitors' interests are modeled and without any modeling of user characteristics or level of knowledge, which require some prior knowledge and more complicated user modeling mechanisms.

Similar approaches have been adopted successfully and described in the literature for museum guides. Kay et al. (2005b), for example, also describes an application of user modeling in a museum, where visitor "interest level" in the content of the exhibits is inferred, recorded and used to customize presentations to visitors. A similar approach was also taken by Hatala et al. (2004), who developed ec(h)o, an "augmented reality interface" for a natural history and science museum. They tracked the visitors in the museum, modeled visitor interaction behavior, and used a domain ontology for inferred interests' representation that later on guided the selection of audio objects suggested to the visitors following a set of inference rules.

In PEACH, the visitor's level of interest in the various concepts and instances is defined on a 5-level scale: "Very Interested", "Interested", "Interested a Little", "Not so Interested" and "Not Interested", as illustrated by Fig. 14.

When a presentation is delivered to the user, newly mentioned concepts and instances are added to the list of interests, and they get a neutral value—"interested a little". Explicit and implicit visitor feedback is used to trigger the inference mechanism that updates the interest values. Explicit user feedback is communicated through the *like-o-meter*, as explained before. Implicit positive feedback is activated by the completion of a presentation delivery to the user, without interruption (e.g., neither extreme negative feedback nor visitor position changed). Explicit feedback has a higher priority than implicit feedback in the sense that

```
1.    <userPosition userId="p02av" device="d004" position="january" orientation="facing" timeStamp="1094473884472" />
2.    <startPresentation id="p02av-1" userId="p02av" device="d004" type="movie" length="125790" location="january" timeStamp= …/>
3.    <startMediaItem id="shot01" userId="p02av" device="d004" type="shot" length="36045" position="january" timeStamp= … >
4.       <concept topic="JANUARY-FRESCO001" />
5.       <perspective type="generic" />
6.    </startMediaItem>
7.    <endMediaItem id="shot01" userId="p02av" device="d004" type="shot" length="36045" position="january" timeStamp= …>
8.       <concept topic="JANUARY-FRESCO001" />
9.       <perspective type="generic" />
10.   </endMediaItem>
11.   <startMediaItem id="shot04" userId="p02av" device="d004" type="shot" length="42000" position="january" timeStamp=…>
12.      <concept topic="SNOWBALLFIGHT-SCENE001" />
13.      <perspective type="generic" />
14.   </startMediaItem>
15.   <wow userId="p02av" userDevice="d004" timeStamp="1094473950166" />
16.   <endMediaItem userId="p02av" type="shot" id="shot04" device="d004" position="january" timeStamp="1094473969354">
17.      <concept topic=" SNOWBALLFIGHT -SCENE001" />
18.      <perspective type="generic" />
19.   </endMediaItem>
20.   <startMediaItem id="shot04" userId="p02av" device="d004" type="shot" length="42000" position="january" timestamp=…>
21.   <concept topic="CASTLE-SCENE001" /><perspective type="generic" /></startMediaItem>
22.   <endMediaItem userId="p02av" type="shot" id="shot04" device="d004" position="january" timestamp=…>
23.      <concept topic="CASTLE-SCENE001" />
24.      <perspective type="generic"  />
25.   </endMediaItem>
…
```

Fig. 15 Example of visitor log following the first segment of scenario 2

explicit feedback is more reliable, so it drives an immediate change in inferred level of interest. Implicit feedback is less reliable; hence it requires accumulation of evidence for every concept (several implicit responses) before changing a visitor's interest level in that given concept.

The inference mechanism starts from the specific concepts mentioned in the presentations and updates their interest value according to the visitors' feedback. Then it follows semantic ontological links to related concepts, thus augmenting the user model with additional, more abstract concepts (see below for a concrete example that illustrates this process). The inference mechanism resembles the *activation network* implemented by prior work in Al Fresco (Stock et al. 1993) and HyperAudio (Sarini and Strapparava 1998), but unlike prior work, where sets of concepts were "activated" or "deactivated", in our case also individual instances are considered, following ontological links in the knowledge base, and the representation itself is much simpler.

An uncertainty factor reflects the fact that there is a need to accumulate evidence about user preferences before changing the inferred level of interest, when dealing with implicit feedback and/or following ontological links to related concepts. Hence, only when the level of certainty reaches a pre-defined threshold, the inferred level of interest the user has in a specific concept is updated.

As a concrete example of how the knowledge-based user model works, let us go on with the second scenario described in Sect. 2.1. Figure 15 presents a partial visitor's log representing the initial part of the second scenario visit (in XML format).

As we can see, the user provides an extreme positive feedback while watching a presentation about the snowball fight scene (appearing as a "wow" event, line 15 in Fig. 15). This explicit feedback is used to trigger the inference mechanism, determining high visitor interest level in the concept SNOWBALL-FIGHT-SCENE001 causing it to be added to the user profile with high interest value (this concept did not appear at the stereotype in Fig. 13 but was added to the user model in Fig. 14). This scene also includes groups of "noblemen" belonging to the "aristocracy", as described in the domain knowledge base (in Fig. 16, bold words illustrate terms used for inference while following the ontological links of the above example).

```
January
 (JANUARY-FRESCO001 :NO-VALUE
  (INSTANCE-OF (FRESCO-ART))
 (HAS-SCENES (SNOWBALL-FIGHT-SCENE001 CASTLE-SCENE001 HUNTING-SCENE001))
  (META-ELEMENTS (BEING0011 DOMINATING001))
  (HAS-THEMES (REPRESENTATION001))
  (HAS-REPRESENTATION (REPRESENTING001))
  (HAS-MAIN-THEME (SNOWBALL-FIGHT001))
  (HAS-MAIN-EVENT (HAVING001))
  (HAS-IMAGES (JANUARY-IMAGE001))
  (HAS-SUBIMAGES (HUNTER-IMAGE001 CASTLE-IMAGE001))
  (HAS-SUBAREAS ())
  (SPECIALTY-OF (JANUARY-MONTH001)))

 (SNOWBALL-FIGHT-SCENE001 :NO-VALUE
  (INSTANCE-OF (SCENE))
  (SCENE-OF (JANUARY-FRESCO001))
 (CONTAINS-ELEMENTS (GROUP001 SNOWBALL-FIGHT001 SNOWBALL001 SNOW001))
  (CONTAINS-EVENTS (HAVING001 THROWING001))
  (PERSPECTIVE (ARISTOCRACY LEISURE NATURE PRESERVATION))
  (META-ELEMENTS ()))

 (SNOWBALL-FIGHT001 :NO-VALUE
  (INSTANCE-OF (SNOWBALL-FIGHT))
  (PARTICIPANTS (ARISTOCRACY))
  (MAIN-THEME-OF (JANUARY-FRESCO001))
  (ELEMENT-OF (SNOWBALL-FIGHT-SCENE001))
  (ACTIVITY-IN (SNOWBALL-FIGHT-SCENE001))
  (HAS-IMAGES (SNOWBALL-FIGHT-IMAGE001)))
```

Fig. 16 Sample portion of the Torre Aquila knowledge base

As a result, the interest levels in the "aristocracy" and "leisure activities" are updated as well. Implicit positive feedback is activated by the completion of a presentation delivery to the user, without interruption (e.g., neither extreme negative feedback nor position changed, line 22 in Fig. 15). In our example, the normal termination of the presentation about the Castle scene, causes the addition of this concept to the user model with an initial value of "interested a little" (Fig. 14).

As a result of the above process, when the user arrives at the February fresco, this knowledge may lead to the presentation of information about the jousting scene, which is another aristocratic activity.

The stored and inferred information is available for use by other components as needed. The presentation composer, while preparing a presentation to be delivered for the user, needs to know about prior presentations delivered, exhibits already visited and the level of interest the visitor has in concepts that are candidates to be included in the next presentation. The report generator needs to know the visit history, details about feedback to specific presentations, their details etc.

An XML-based reporting and querying language was developed for this purpose. XML messages are sent by the system components reporting events, and queries are posed by these components and answered by the User Model as needed. The queries may address facts such as whether the visitor has seen a specific shot or been at a specific position. They also may inquire about the level of interest the visitor has in a specific concept or request a summary of the whole visit up to the point of time when the report is requested. The User Model searches the events log and responds to these queries regarding specific events, or it reports about the level of interest the visitor has in a specific concept.

5.3 Template-based propagation of interests

As an alternative to the knowledge-based UM approach presented above, we implemented a prototype based on the propagation of interests into a network of content (Rocchi and Zancanaro 2003). This component requires less work at the knowledge representation level,

and is preferable for fast prototype realization of the multimedia guide described in Sect. 3.3. This model works over the network of video templates a described in Sect. 3.3.1 (see sample network in Fig. 10).

Interest shown towards a presentation built from a video template may be propagated to nearby templates. For instance, if a template describes the overall fresco and the visitor expresses interest using the *like-o-meter* widget, the system then propagates interest to templates which are connected to the current one. In interactions that follow, this information is more likely to be selected for presentation.

The interest value assigned to each template ranges from -2, meaning current lack of interest, to $+2$, meaning strong interest. As feedback is received through the interface, the system updates the interest value of the current template and also propagates such information to its connected templates.

As explained in Sect. 3.3.1, there are several types of presentations and the type of presentation is relevant when inferring the interest of the visitor and the resulting action to be taken. Propagation is based on the following dependency relations:

1. *Introduction affects Abstract*: a positive or negative degree of interest, expressed when the system presents the general introduction to an exhibit, is propagated to all the abstracts pertaining to the same exhibit (the abstract is the basic information on an exhibit with respect to a certain topic).
2. *Abstract affects Content*: the degree of interest towards the abstract updates the value of its related content, (i.e. a more detailed description of the exhibit with respect to the abstract). When a content template has an interest value greater than zero, it is selected and presented to the visitor right after the abstract.
3. *Content affects Follow-up*: the degree of interest towards contents affects its connected follow-ups (if any). Follow-ups are selected and proposed to the visitor when they have an interest value of $+2$. A follow-up contains general information about a topic shared by two or more exhibits.
4. *Follow-up affects Content*: the degree of interest towards follow-up is propagated to all of the connected contents. Follow-ups act as bridges by propagating the interest values on one exhibit to other similar exhibits.

The advantage of propagating interests directly over the video template network rather than reasoning on the semantic links of a knowledge base is that the latter is difficult to build and to update. The disadvantage of this variant is that it allows less flexibility than the system based on explicit user modeling. In addition, the representation cannot be used easily for unconnected content (e.g., a temporary exhibition on a similar theme).

6 Underlying distributed architecture and user tracking

In this section, we provide a short overview of aspects of the technical infrastructure used in PEACH which are relevant for this paper. Specifically, we focus on the architecture, which has determined how the system has been developed, and on the visitor tracking functionality, which is critical to support system adaptivity and user modeling of mobile users.

6.1 Multiagent architecture

As a complex set of software components and interfaces, the PEACH museum system has been distributed over a number of devices, including user PDA's, static objects (e.g., location

sensors, plasma screens), and back end computers. In an active museum environment, many autonomous components exist and operate in parallel and simultaneously seek and provide services in a highly distributed and dynamic manner. To overcome limitations of the traditional client-server approach (used in many of the other similar applications surveyed in Sect. 1.1), an architecture was designed according to an innovative multi-agent approach based on group communication (Busetta et al. 2002). In order to be flexible and support the evolution of a user's context (e.g., the visitor's location and state and locally available services), the architecture was defined in terms of abstract *roles* played by agents and the *policies* they need to follow when delivering a service.

We call a set of agents playing a certain role according to a given policy an *implicit organization* (Busetta et al. 2003). The agents within an implicit organization may be very different from each other and may run anywhere (e.g., in the *Positioning* implicit organization, agents are represented by different localization subsystems exploiting different hardware devices and interpretation processes). The main advantage of this approach is that, by properly designing a system, each function becomes a distributed system in itself that easily adapts to changes in available resources (e.g., PDA's coming and going, on-the-fly addition of new devices, device failures, upgrades of software while the system is running, and so on) and context and conditions (e.g., system workload, user location, user preferences, and so on).

The requirement analysis and architectural design phases led us to identify a few implicit organizations, each corresponding to some specific system functionality. The *Positioning* implicit organization is responsible for reporting the users' current positions periodically and/or by request. The *Spatial information broker* provides positioning relations between visitors and museum exhibits. The *User modeler* implicit organization has been described in the previous section. The *Information broker*, the *Presentation composer*, the *Presentation client* (which provides a variety of presentation capabilities according to location and available equipment—on static or mobile devices, with / without audio, on screens of different sizes or orientations, and so on), and the *Report Generator* are other examples of implicit organizations.

The architecture defines the relationships among organizations as a set of scenarios described by a context, a triggering event, and the protocol that should ensue. For instance, the delivery of a presentation may be triggered by the Spatial Information Broker listening to a report by the Positioning agent and determining that the user is in close proximity to a certain exhibit. The protocol that follows involves the Presentation Composer interacting with the User Model, the Information Broker and the Presentation Client in order to (1) get knowledge on the user and the exhibit, (2) find out the features of the best suited output device at the user's location, and finally (3) prepare the presentation.

Internally, each implicit organization adopts the policy that is best suited to the service it has to deliver in the current context. For instance, the *Presentation Composer* adopts a competitive policy to select which one is "best" at preparing a specific presentation within a pool of agents of different capabilities and computational resources. The competition boils down to either a bidding system (i.e., an auction) or a simple race (the first to answer wins), depending on whether priority is given to quality of the output or to speed of delivery, which in turn depends on the current system workload. By contrast, the *Information Broker* uses a collaborative policy, such as merging the data fetched by each agent, in order to return the greatest possible amount of information to queries. In the first version, in which only individual visitors are supported, the *Presentation Client* adopts a competitive policy based on proximity, quality, and user preferences about output devices; this policy may change in a future version, for instance with presentations that will be delivered to a group of people simultaneously.

The infrastructure supporting implicit organizations is based on agent group communication, i.e. the ability of sending messages to an unlimited number of agents that have shown a common interest. To this end, we have defined a conceptual framework that we call *channeled multicast,* and implemented it in our own experimental framework, called LoudVoice (Busetta et al. 2002). The idea is to let agents create or "tune" to any number of channels where anybody can send a message to anybody else and listen to everybody's messages, no matter their intended destinations, in a way vaguely similar to chat rooms. Messages can be addressed either to specific listeners or to any agent that matches an expression in an application-specific language.

A small scale crash test was performed to measure the robustness of the architecture. The configuration used for the test encompassed an access point (3COM 11 g), 4 PDA's and 4 laptops: 2 Pentium 4 2.3 GHz, one AMD 3000+ and one Pentium 4 3 GHz. All the laptops had 512 M of RAM and an integrated network card (respectively, two with a 11 Mb card and two with a 54 Mb card). The system was divided on the four laptops as follows: one laptop ran a MySQL database, the servlet engine Tomcat and the *Information Broker* agent; another laptop ran all the other agents. The other two laptops ran 200 simulations of users who randomly moved around (one movement every 30 s) and retrieved information. Similar simulations were also run on the four PDA's: each PDA ran a single interface complete with the graphical display and an agent that simulates movements and user actions.

During the test, the network load never exceeded 25% of the available bandwitdh. The laptops that ran the architecture worked with the CPU fully loaded but without unsteadiness or crashes for several hours. Moreover, the CPU load reduced after an initial peak. The available memory was barely sufficient for the laptop that ran MySQL and Tomcat (the two applications should run on two different machines). No problems were detected on the laptop that ran the remaining agents.

6.2 Visitor position tracking

In order to provide location-aware presentations and to support the user model (we have discussed earlier the information coming from the position choice of the visitor and the time spent in front of an exhibit) the architecture has to support tracking of a visitor's position both at a low-level (i.e. by filtering and fusing sensors' signals) and at a higher-level of reasoning (i.e., by calculating proximity to exhibits and presence into wider areas such as rooms). The PEACH architecture exploits the mechanism of channeled multicast to tackle the first task and provides an initial infrastructure to link exhibits to areas for the second task.

In the course of the PEACH project, we have experimented with several types of sensors: Infrared beacons (IR), Radio Frequency Identification (RFID) transponders and transceivers, technologies based on Wi-fi signal strength, and accelerometers.

Infrared beacons can send a unique code to a distance of up to eleven meters. Since the IrDA standard is used,[1] any infrared-equipped device can receive it. The code can be used to identify a position in the space very effectively and robustly. The beacons are precise, can be powered by standard batteries and have very low power consumption. Several beacons with different sending ranges, all installed in the same location, allow the system to roughly distinguish the distance of the mobile device to that location. However, they have several drawbacks, requiring line of sight and the visitor has to deliberately point the device toward the beacon.

---

[1] Standard protocols for infrared wireless communication are defined by the Infrared Data Association (IrDA). http://www.irda.org/

RFID technology overcomes some of these limitations. In PEACH we have experimented with active tags where the PDA works as a transponder. RFID tags can be detected from 1 to 100 meters and do not require line of sight. Yet, they cannot provide any information about the orientation of the visitor, but only their presence in the proximity of an area.

Technologies based on Wi-Fi signal strength allow continuous tracking of the visitor and therefore, by using temporal information, it is possible to infer some degree of orientation as well. Yet, in our experience they have proven to be lacking in precision and quite unreliable for museum applications, in particular because usually the tracking has to be performed in rather small spaces. This kind of technology, moreover, may be very expensive since an overabundance of access points is desirable in order to increase localization precision.

Finally, accelerometers can be used to provide more reliable information about the orientation of the device in 3 dimensions. This allowed us to (a) estimate the orientation of the visitor, and (b) determine whether the user is looking at the screen of the device (i.e., when the device is being held within a certain vertical range).

Although no single one of these technologies is perfectly suited for our needs, their combination can provide the right level of accuracy and robustness required for a location-aware delivery of personalized information. The PEACH architecture provides a *Position channel*, devoted to the communication of sensor information. The Spatial Information Broker monitors the channel and is in charge of fusing the incoming sensorial information to determine the actual position of each visitor. The computed position is then re-posted on the same channel, so that all the other agents know both the filtered information and all the unfiltered data.

Moreover, the broker abstracts the low level tracking information and provides position information in terms of the area in which the visitor is located. An area can be a floor, a room or a region of a room. Areas can be nested on the basis of semantic relevance for the visitors or just for positioning, when a region is not completely covered by sensors.

So, for example, Wi-fi-based technologies can be used to detect the coarse-grained information of visitors entering and leaving rooms, while RFID's can detect the position inside the rooms or the proximity to isolated exhibits. Finally, IR can be used when several exhibits are placed very closely to each other and the visitor explicitly "points" at the exhibit she is interested in.

Areas are associated with exhibits using one-to-many relations and the association is kept in the multimedia database. This permits distributing amongst different components the responsibility of organizing the space and mapping the content into the space (exploiting the distributed architecture).

The "captured" IR signals do not tell us, however, which exact exhibit detail the user is currently observing. Even if it is beyond the scope of the specific short examples discussed in this paper, let us quickly describe another development of PEACH (only mentioned in Sect. 2). The question of how to detect precisely the focus of attention of the visitor is not trivial, if we do not want to bring in goggles, helmets or other devices that would interfere with the visitor's freedom and their normal behavior. To solve the problem we have investigated a paradigm of interaction based on the point-and-shoot use of modern digital cameras (see Albertini et al. 2005). The visitor's PDA is equipped with a webcam and artificial vision technology is employed to recognize which painting or detail the visitor is pointing at. In order to ask for information about a painting or a detail, the visitor just points the webcam toward the painting. A copy of the camera view is displayed on the PDA screen, while color-reduced still images are sent to the vision-recognition engine. If the engine recognizes a scene, an identifier is sent back to the interface that queries the system for a presentation. If a multimedia presentation for that scene is available the system prompts its title, in the

form of a textual label displayed on the PDA screen and if the user selects it, a multimodal presentation is played.

The visual engine employed is a modified version of the content based image retrieval system COMPASS (Andreatta et al. 2005). It classifies an unknown query image by comparing it to a database of images containing relevant details of all the exhibits (many images are stored in the database for each detail in order to provide views on many different perspectives). For each image, a similarity score is computed and the exhibit (or the detail) that corresponds to the highest scored image is selected.

7 Experiments with users

Though there has been no evaluation of a fully integrated system, we conducted a number of user studies to evaluate the ideas underlying the many PEACH components.

A "layered approach" was proposed by Totterdell and Boyle (1990) in order to break down adaptation in its constituents and to separately evaluate each constituent. The success of adaptation can be addressed at two distinct layers, reflecting the main phases of the adaptation processes, namely user modeling and adaptive output (e.g., multimodal presentations or visit reports). Although the two phases are both important for the success of adaptation, they are also independent because the same user modeling outcomes may result in significantly different adaptation decisions. Hence, separate evaluations can contribute to the generalization and re-use of the results across different applications. This separation can be seen as one peculiarity that differentiates the evaluation of user adaptive systems from the evaluation of "non-adaptive" systems (Gena 2005). The evaluations we performed were mainly addressed at the presentation and interaction layer.

7.1 A summary

The evaluation of the idea of using cinematography as a paradigm for presentation of information was initially investigated in (Nardon et al. 2002) while the assessment of this approach in a mobile setting was discussed in (Alfaro et al. 2004). A heuristic evaluation of the automatically generated video clips with two film directors was discussed in Callaway et al. (2005a).

A specific study aimed at investigating the possible interferences caused by the PDA during a museum visit. Our main hypothesis was that using cinematic presentation techniques would ensure greater cohesion, which in turn would bring to a higher level of predictability and a less stressful coordination of attention between the PDA and the exhibits. The experimental data supported our hypothesis since the number of eye movements from the PDA to the fresco and vice versa were significantly fewer when the cinematic techniques were applied with respect to a control condition (a standard multimedia presentation).

Evaluation of aspects specifically concerning natural language generation is usually devoted to whether adaptivity is appropriately taken into account. While the report generator has not been evaluated thoroughly, a small study on the acceptance of personalized report generation was presented in (Goren-Bar and Prete 2005). That preliminary study aimed at understanding visitors' attitude towards reports written with different degrees of personalization and different presentational strategies (sequential vs. thematic reports). The resulting findings seem to suggest that personalization in reports might strengthen the perception of "being watched" during the visit; yet, we have to say that the practical advantage of personalized reports is difficult to perceive in the case of a small exhibition. In any event, to counter that impression

we could allow visitors to be active and contribute during the visit to the report generation, for instance by enabling them to "save" during the tour some images they prefer and, at the end, have a report including a picture on those works of art they had previously chosen.

Within the scope of the PEACH project a medium scale user study (42 participants) has been conducted evaluating the effectiveness of a lifelike character guiding users' attention during multi-device presentations (i.e. presentations spanning over two or more different presentation devices such as public displays, spatial audio systems or personal digital assistants) from one device to another (see Kruppa and Aslan 2005). This method was compared against one in which an animated symbol instead of the character was used to guide the users' attention as well as against a control in which the users' attention was not at all actively guided during presentations. Participants of the experiment were confronted with a novel presentation technique combining both public and personal presentation systems which used both types of devices in parallel in order to allow for personalized presentations tailored to the interests of each individual user. The system used a public display with an associated public audio system as the main presentation means, and in addition used the personal device of each participant from time to time to present personalized, individual presentations in parallel to the public presentation. Even though objective measures did not reveal any significant difference in recall performance between the different focus guidance methods, subjective results clearly indicated that the participants had a high preference for the lifelike character (57% preferred the character, 24% preferred not to be guided actively, 7% preferred the animated symbol and 12% had no preference at all).

The feeling of being in control over the interaction is an important component of people's comfort with technologies (Norman 1998); several studies have discussed these issues in the context of adaptive systems (see for example Jameson and Schwarzkopf 2002), all of them pointing to the opportunity of increasing the controllability of all aspects of system adaptation. In an attitudinal study (Goren-Bar et al. 2006), we showed how personality traits correlate with the degree of acceptance of personalization and how the most relevant traits are specifically those relating to the notion of control, in particular, *Conscientiousness* and *Emotional Stability* captured by the Big Five model (McCrae and John 1992) and the *Locus of Control* (LoC, Rotter 1966).

The Big Five model posits five dimensions as relevant factors in assessing an individual personality. The dimension of *Conscientiousness* is concerned with the way we control, regulate and direct our impulses. People high on conscientiousness have a good control on their own impulses, are confident in their ability to accomplish things, are well organized, have a clear sense of their duties and obligations and do not easily become discouraged. With respect to adaptivity, according to our data, these subjects tend to dislike delegating the start of a presentation to the system on the basis of location-awareness, and dislike content adaptation with respect to the user model. *Emotional Stability* refers to the way people react to stressful situations: emotionally stable people are at ease most of the time, usually calm and not easily upset. According to our data, emotionally stable subjects have a positive attitude towards content adaptation with respect to the user model. The other three dimensions, *Extraversion* (the quantity and intensity of one's interpersonal reactions, sociability and emotional expressiveness), *Agreeableness* (which reflects the individual differences with respect to cooperation and social harmony) and *Creativity* (consists of openness to, and proactive seeking and appreciation of new experiences) do not seem to play a role in the acceptance of adaptivity.

The Locus of Control is sensitive to where people feel the control over one's beliefs and actions is located. When internal, the subject feels that she is the controller; an external LoC signals that the subject's beliefs and actions are substantially independent of her own

influence. According to our data, people with Internal LoC have a general positive attitude towards adaptivity and they do not seem to be afraid to delegate some control, probably because they feel confident that they can always recover it. The subjects with a more external orientation, on the other hand, seem more interested in maintaining a direct control on the device, and praised nonadaptivity.

As said above, this attitudinal study was aimed at investigating the attitudes and dispositions of (potential) users towards the very idea of adaptivity, rather than focusing on the usability of specific systems.

User control is not the only critical issue for adaptive systems and intelligent user interfaces in general. As a new paradigm of interaction with computers, intelligent interfaces often violate further major usability principles, as pointed out by Höök (2000), such as: make the behavior of the system predictable (in the sense that the same input always causes the same response), and make the system transparent so that the user can understand its inner working. Therefore special caution must be given to the definition of appropriate studies. Gena (2005) discusses several methodologies for the evaluation and the user-centered design of adaptive systems. In our own work, we always avoided a direct comparison of an adaptive guide with standard audio-guides normally found in museums because such a comparison may be biased by the limited functionalities of the prototypes (including insufficient coverage). We rather preferred to investigate the attitude of the visitors toward adaptation, as in the study briefly described above, or to analyse the actual users' behavior as in the study presented below in Sect. 7.2. In the spirit of the "layered approach" proposed by Totterdell and Boyle (1990), we are also conducting a study to derive an empirically validated model for the acceptance of adaptive multimedia guides in a museum setting, by considering both extrinsic (e.g., ease of use, perceived usefulness) and intrinsic (involvement) motivational factors.

7.2 Assessing the adaptive multimedia guide based on like-o-meter

We performed an initial assessment of the acceptability of the *like-o-meter* interface for control and delegation of the adaptation process. This study, as opposed to the mock-up-based studies for the design of the interface described in Sect. 3.4, is based on the use of the working system in the ecological setting.

For this experiment, the system encompassed the use of the engine for content adaptation (see Sects. 3.3 and 5.3) but not the virtual characters with the associated stereotypic approach. Of course the amount of difference in the presentations could be larger in the latter case, yet for this evaluation we were interested instead in investigating the subtler and more dynamically tailored characteristics of the propagation-based technique.

One hundred and forty-three actual public visitors of the Torre Aquila were invited to test the adaptive multimedia guide based on the *like-o-meter* described above. Sixty one visitors were males and 82 females. Their age ranged from 20 to 79 years (mean = 47, std.dev = 15.87). They were recruited at the entrance of the Museum of the Buonconsiglio and received a free ticket for visiting the castle as a reward for participating in the experiment. The data of 33 visitors were discarded because they did not completely fill in the questionnaires. The data analyzed below is then based on the questionnaires and the observations on the remaining 110 visitors (49 males and 61 females with an average age of 45 year old—median = 47, std. dev = 15.534).

Two questionnaires were used in the study. The first questionnaire was aimed at assessing the dimensions of attitude toward art (4 items; Stokmans 1999) and attitude toward technology (6 items; Popovich et al. 1987). The second questionnaire was aimed at assessing their experience in using the multimedia guide along diverse dimensions (all as Likert scales).

Table 11 Average scores for attitude toward technology on the four studied dimensions: Involvement, Control, Ease and Intention to use

	Global_Involvment	Global_Control	Easiness	IntToUse
Positive	−0.46	−0.31	−0.32	−0.32
Negative	0.65	0.43	0.44	0.44

The dimensions investigated were measured on many different scales and they were then combined and factorized in the following four dimensions:

- Control: obtained by combining items from perceived quality of information, perceived spatial orientation, and control (Novak et al. 2000) with a combined reliability of 0.797.
- Involvement: obtained by combining items from flow, involvement, perceived time distortion and perceived presence (Novak et al. 2000; Csikszentmihalyi 1990) with a combined reliability of 0.904.
- Ease: obtained by combining items from ease to understand, ease of use and clarity of feedback (Davis 1989) with a combined reliability of 0.890
- Intention to Use: obtained by combining items from the Technology Acceptance Model (Davis 1989) with a combined reliability of 0.800.

Additionally, two four-choice questions were aimed at assessing the degree of understanding of the interface and the use of the like-o-meters ("when you are interested/not-interested in a topic presented by the guide, what did you do: a. you pressed the smiling face; b. you moved to another fresco; c. you pressed the sad face; d. you did nothing") combined into a binary measure called Correct Use (true if the users select answers a and c respectively, false otherwise).

The study was conceived as an exploratory study on the usage of an adaptive guide. We chose to use age, sex and attitude toward technology as factors. Age and attitude toward technology have been transformed into two level factors. The age groups were obtained by splitting the ages according to the median value—younger visitors with age less than or equal to 47 and older visitors with age of 48 or higher. Attitude toward technology was considered positive when scored above the median and negative when scored less.

All the variables were standardized on distributions with mean 0 and standard deviation 1. Higher scores indicate disagreement. Distributions show clear left asymmetries, indicating a general agreement on all the dimensions: that is, our visitors generally felt in control of the system, had good involvement, found it quite easy to use and had a propensity to use it again in the future.

In particular, regarding age group, significant differences emerged with respect to involvement and intention to use. Young visitors' average score for intention to use is 0.24 while for older visitors is −0.23. Therefore, the latter have more propensities towards future use of the system.

Average scores for attitude toward technology are shown in Table 11. As expected, a positive attitude toward technologies leads to greater agreement on each dimension.

Regarding the variable Correct Use (the binary measure that indicates whether the subjects understood how the *like-o-meter* works), we wanted to test whether there is a relationship among this variable and the variables used as factors in the study, namely age group, sex and attitude toward technology. We used a fitted loglinear model using backward elimination. The best model comprises *AgeGroup*CorrectUse + AgeGroup*AttitudeTowardTech* which means that there is a relationship between age and the correct use of the system as well as between age and attitude toward technology. Namely, young visitors use the system in a

Fig. 17 Total number of
expressions of interest (both
positive and negative)
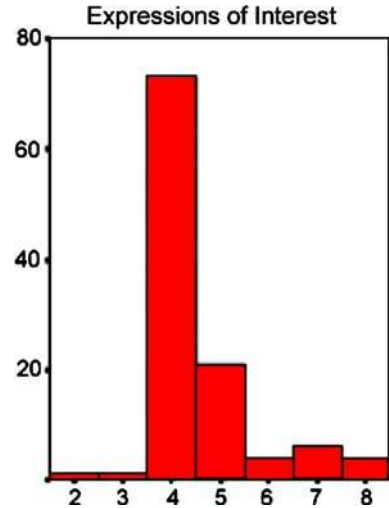
correct way more often and they more often have a positive attitude toward technology with
respect to older visitors.

As far as the actual use of the *like-o-meter* is concerned, the total number of expressions
of interest (both positive and negative) ranges from a minimum of 2 to a maximum of 8 for
each visitor (average 4.5, median 4 and standard deviation 1.08, see Fig. 17).

We used K-means to divide the visitors into those who express their interest often and those
who express their interest less often. The group with high expression of interest consists of 35
visitors with an average of 5.8 expressions (S.E. = 0.186) while the group with low expression
of interest contains 75 visitors with an average of 3.96 expressions (S.E = 0.02). Running
an ANOVA using the clusters as a factor and the variables above as dependent variables,
significant differences were found only for *Age* and *CorrectUse* (respectively, $F = 5.99$,
sig. < 0.01 and $F = 609.955$, sig. < 0.01): the group that expresses more interest tends to
have an older age (50.23 vs. 42.61) but also tends to display a less correct use of the system.
Very interesting is the variable *GlobalControl* which is close to significance ($F = 3.257$;
$p = 0.074$): those who express more interest feel more in control of the system (0.25 vs.
-0.11) even if, as noted above, its usage is not correct.

In summary: the participants generally felt in control of the system, had good involvement,
found it quite easy to use and expressed their intention to use it again in the future. We found
that a positive attitude toward technology is related to a greater agreement on each dimension
(control, involvement, ease to use, and intention to use in the future), meaning that probably
this attitude enabled the visitors to be more open to the experience this new technology has to
offer during the visit at the museum. The surprising results were that although young visitors
in our sample had generally a more positive attitude toward technology than older visitors
and they used the system in a correct way more often (meaning that they displayed better
understanding of the *like-o-meter*) the older visitors expressed a greater tendency toward
their intention to use the guide in the future.

We might speculate that since older visitors are less acquainted with new technologies
and less skilled to properly operate new systems, they were surprised by the good experience
they felt while using it. Their tendency towards using the guide in the future is in fact a very
encouraging result, indicating that this type of technology can be used by the museum public

in general and not especially by young people with technological background as might be expected.

8 Conclusions

We have described an advanced and integrated framework for museum visits, focusing on aspects where adaptivity is central. The creation of a complex and multifaceted system to serve the visitor required conducting research on various intelligent technologies and assembling them into a coherent picture.

In this paper we first set out a scenario and showed the diverging behavior of the system in two exemplary, apparently similar visits, in order to illustrate the flexibility in our adaptive system. We then described various components that contribute to multimedia mobile presentations by integrating language and images, and the automatic generation of video clips. We have specifically focused on adaptive documentaries for the mobile visitor. We have also discussed the added value of a virtual character that plays the role of presenter and helps bootstrap the user model, and the seamless integration of stationary displays and mobile systems, all of which are context-aware and adapted to the visitor. We have discussed interface design issues and in particular effective ways for providing feedback to the system, influencing future system behavior in the course of the visit. We have also introduced technology that intervenes at the end of the visit, which is concerned with producing a personalized report of the specific visit and emphasizing elements that seemed of major interest to the visitor. The report becomes an entry point for acquiring more information on the themes that appeared of interest for the visitor and for attracting her to subsequent visits.

The basic underlying user modeling and adaptation mechanisms that underpin the behavior in the examples have been briefly described. A novel multiagent architecture for hosting and using the above mentioned technologies has been introduced providing a framework for expansion in a highly dynamic museum environment. For an essentially mobile adaptive system, positioning is a fundamental aspect; and we have presented various solutions that have been introduced and experimented with within PEACH. Experiments with real visitors have provided empirical validation for a number of technologies introduced in PEACH, some of which have been reported in the final section of this paper. Yet an overall evaluation of the entire integrated framework would be hard to perform, at least for the time being. This constitutes a possible experimental research topic for the future, and will require an appropriate methodology and the possibility of accumulating and comparing results over time. In fact the overall goal for the introduction of technology in the cultural heritage context is to raise the level of interest and motivation of visitors; this can amount to making visitors come again to the site and/or visit other sites with increasing competence and pleasure, while transmitting this interest to others. Other issues related to engineering and portability of a complex system deriving from this research, such as maintainability and appropriate authoring interfaces, are also topics not addressed yet, beyond expressions of guidelines (e.g., Katz et al. 2006).

Overall, PEACH has helped define a potentially very rich and new class of ambient systems at the service of mobile visitors. Many research topics are open to further investigation. Among those we are currently focusing our research on, and that we would like to emphasize, are importing a user model from past experiences, and the theme of group visits: introducing technology that helps the group experience to be more effective, pleasurable, memorable and interesting than the collection of individual experiences.

References

Abowd, G., Atkeson, C., Hong, J., Long, S., Kooper, R., Pinkerton, M.: Cyberguide: a mobile context-aware tour guide. Wirel. Netw. **3**(5), 421–433 (1997)

Albertini, A., Brunelli, R., Stock, O., Zancanaro, M.: Communicating user's focus of attention by image processing as input for a mobile museum guide. The 10th International Conference on Intelligent User Interfaces, pp. 299–301. San Diego, CA (2005)

Alfaro, I., Nardon, M., Pianesi, F., Stock, O., Zancanaro, M.: Using cinematic techniques on mobile devices for cultural tourism. Inform. Technol. Tourism **7**(2), 61–71 (2004)

Andreatta, C., Lecca, M., Messelodi, S.: Memory-based object recognition in digital images. The 10th International Fall Workshop – Vision, Modeling, and Visualization, pp. 33–40. Erlangen, Germany (2005)

Aoki, P.M., Grinter, R.E., Hurst, A., Szymanski, M.H., Thornton, J.D., Woodruff, A.: Sotto Voce: exploring the interplay of conversation and mobile audio spaces. In: Wixon, D. (ed.) The SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, pp. 431–438. ACM Press, New York (2002)

Arijon, D.: Grammar of the Film Language. Communication Arts Books, Hasting House, New York (1976)

Bares, W., Grégoire, J., Lester, J.: Realtime constraint-based cinematography for complex interactive 3d worlds. The Tenth National Conference on Innovative Applications of Artificial Intelligence, pp. 1101–1106. Madison, Wisconsin, USA (1998)

Baus, J., Cheverst, K., Kray, C.: A survey of map-based mobile guides. In: Meng. L., Zipf, A., Reichenbacher T. (eds.) Map-based Mobile Services – Theories, Methods and Implementations, pp. 197–213. Springer, Heidelberg (2005)

Bellotti, F., Berta, R., de Gloria, A., Margarone, M.: User testing a hypermedia tour guide. IEEE Pervasive Comput. **1**(2), 33–41 (2002)

Benelli, G., Bianchi, A., Marti, P., Not, E., Sennati, D.: HIPS: hypen-interaction within physical space. International Conference on Multimedia Computing and Systems, pp. 1075–1078. Florence, Italy (1999)

Berkovich, M., Date, J., Keeler, R., Louw, M., O'Toole, M.: Discovery point: enhancing the museum experience with technology. Conference on Human Factors in Computing Systems, pp. 994–995. Florida, USA (2003)

Berkovsky, S., Gorfinkel, A., Kuflik, T., Manevitz, L.: Case-based to content-based user model mediation. Workshop on Ubiquitous User Modeling, held in conjunction with ECAI-2006. Riva delGarda, Italy (2006)

Bohnenberger, T., Jacobs, O., Jameson, A., Aslan, I.: Decision-theoretic planning meets user requirements: enhancements and studies of an intelligent shopping guide. Pervasive computing: Third International Conference, pp. 279–196. Munich, Germany (2005)

Busetta, P., Donà, A., Nori, M.: Channeled multicast for group communications. The First International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 1280–1287. Bologna, Italy (2002)

Busetta, P., Merzi, M., Rossi, S., Legras, F.: Intra-role coordination using group communication: a preliminary report. International Workshop on Agent Communication Languages and Conversation Policies, pp. 231–253. Melbourne, Australia (2003)

Butz, A.: Anymation with CATHI. The Fourteenth National Conference on Artificial Intelligence and Ninth Innovative Applications of Artificial Intelligence Conference, pp. 957–962. Providence, Rhode Island, USA (1997)

Callaway, C., Lester, J.: Narrative Prose Generation. Artif. Intell. **139**(2), 213–252 (2002)

Callaway, C., Not, E., Novello, A., Rocchi, C., Stock, O., Zancanaro, M.: Automatic Cinematography and Multilingual NLG for Generating Video Documentaries. Artif. Intell. **165**(1), 57–89 (2005a)

Callaway, C., Kuflik, T., Not, E., Novello, A., Stock, O., Zancanaro, M.: Personal reporting of a museum visit as an entrypoint to future cultural experience. The 10th International Conference on Intelligent User Interfaces, pp. 275–277. San Diego, CA (2005b)

Cheverst, K., Davies, N., Mitchell, K., Friday, A., Efstratiou, C.: Developing a context-aware electronic tourist guide: some issues and experiences. The CHI 2000 Conference on Human factors in computing systems, pp. 17–24. The Hague, Netherlands (2000)

Cheverst, K., Mitchell, K., Davies, N.: Design of an object model for a context sensitive tourist guide. Comput. Graph. **23**(6), 883–891 (1999)

Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper and Row, New York (1990)

Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart. **13**, 319–340 (1989)

Dey, A., Salber, D., Abowd, G.: A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. Hum-Comput. Interact. J. **16**(2–4), 97–166 (2001)

Efstratiou, C., Friday, A., Davies, N., Cheverst, K.: A platform supporting coordinated adaptation in mobile systems. The 4th IEEE Workshop on Mobile Computing Systems and Applications, pp. 128–137. Callicoon, New York (2003)

Friedman, D.A., Feldman, Y.A.: Knowledge-based cinematography and its applications. The 16th European Conference on Artificial Intelligence, pp. 256–262. Valencia, Spain (2004)

Gena, C.: Methods and techniques for the evaluation of user-adaptive systems. Knowledge Eng. Rev. **20**(1), 1–37 (2005)

Gena, C., Torre, I.: The importance of adaptivity to provide on-board services. A preliminary evaluation of an adaptive tourist information service on board vehicles. Appl. Artif. Intell. **18**(6), 549–580 (2004)

Goren-Bar, D., Graziola, I., Pianesi F., Zancanaro M.: The influence of personality factors on visitor attitudes towards adaptivity dimensions for mobile museum guides. User Model. User-Adapt. Interact. **16**(1), 31–62 (2006)

Goren-Bar, D., Graziola, I., Rocchi, C., Pianesi, F., Stock, O., Zancanaro, M.: Designing and redesigning an affective interface for an adaptive museum guide. In: Tao J., Tan T., Picard R.(eds.) Affective Computing and Intelligent Interaction ACII-2005 First International Conference on Affective Computing and Intelligent Interaction. Beijing, China (2005)

Goren-Bar, D., Prete, M.: Report on a museum tour report. In: Maybury M., Stock O., Wahlster W. (eds.) INTETAIN 2005: First Conference on Intelligent Technologies for Interactive Entertainment, pp. 230–234. Springer, LNAI 3814 (2005)

Grinter, E., Aoki, P. M., Hurst, A., Szymanski, M.H., Thornton, J.D., Woodruff, A.: Revisiting the visit: Understanding how technology can shape the museum visit. ACM Conference on Computer Supported Cooperative Work, pp. 146–155. New Orleans, LA (2002)

Halliday, M.A.K., Hasan, R.: Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective. Deakin University Press. (1985)

Halper, N., Oliver, P.: CamPlan: A camera planning agent. The AAAI Spring Symposium Workshop on Smart Graphics, pp. 92–100. Stanford, USA (2000)

Hatala, M., Kalantari, L., Wakkary, R., Newby, K.: Ontology and rule based retrieval of sound objects in augmented audio reality system for museum visitors. The 2004 ACM Symposium on Applied computing, pp. 1045–1050. Nicosia, Cyprus (2004)

Heckmann, D., Schwartz, T., Brandherm, B., Schmitz, M.: von Wilamowitz-Moellendorff, M.: GUMO – the general user model ontology. The 10th International Conference on User Modelling, pp. 428–432. Edinburgh, UK (2005)

Höök, K.: Steps to take before IUIs become real. J. Interact. Comput. **12**(4), 409–426 (2000)

Jameson, A., Schwarzkopf, E.: Pros and cons of controllability: An empirical study. In: Brusilovsky P., Conejo E. (eds.) Adaptive Hypermedia and Adaptive Web-based Systems: AH2002, pp. 193–202. Springer Verlag, Heidelberg, (2002)

Karp, P., Feiner, S.: Automated presentation planning of animation using task decomposition with heuristic reasoning. Graphics Interface'93, pp. 118–127. Toronto, Ontario Canada (1993)

Katz, S., Kahanov, Y., Kashtan, N., Kuflik T., Graziola, I., Rocchi, C., Stock, O., Zancanaro, M.: Preparing personalized multimedia presentations for a mobile museum visitors guide – a methodological approach. In: Trant, J., Bearman, D. (eds.) Museums and the Web 2006, Archives & Museum Informatics, Toronto: published March 1, 2006 at http://www.archimuse.com/mw2006/papers/katz/katz.html. (2006)

Kay, J., Kummerfeld, B., Lauder, P.: Consistent modelling of users, devices and sensors in a ubiquitous computing environment. User Model. User-Adapt. Interact. **15**(3–4), 197–234 (2005a)

Kay, J., Lum, A., Niu, W.: A scrutable museum tour guide system. The 2nd Workshop on Multi-User and Ubiquitous User Interfaces, Sonderforschungsbereich 378, Saarland, pp. 19–20 (2005b)

Krüger, A., Butz, A., Müller, C., Stahl, C., Wasinger, R., Steinberg, K., Dirschl, A.: The connected user interface: realizing a personal situated navigation service. The 9th International Conference on Intelligent User Interfaces, pp. 161–168. Madeira, Portugal (2004)

Krüger, A., Kruppa, M., Müller, C., Wasinger, R.: Readapting multimodal presentations to heterogenous user groups. AAAI-Workshop on Intelligent and Situation-Aware Media and Presentations, T. R. WS-02–08, AAAI Press, pp. 46–54. (2002)

Kruppa, M., Aslan, I.: Parallel presentations for heterogeneous user groups – an initial user study. In: Maybury, M., Stock O., Wahlster, W. (eds.) INTETAIN 2005: First Conference on Intelligent Technologies for Interactive Entertainment, pp. 54–63. Springer, LNAI 3814 (2005)

Kruppa, M., Krüger, A.: Concepts for a combined use of personal digitalassistants and large remote displays. Simulation und Visualisierung, pp. 349–361. Magdeburg, Germany (2003)

Long, S., Aust, D., Abowd, G., Atkeson, C.: Cyberguide: Prototyping context-aware mobile applications. Conference on Human Factors in Computing Systems, pp. 293–294. Vancouver, Canada (1996)

Malaka, R., Zipf, A.: DEEP MAP – Challenging IT research in the framework of a tourist information System. In: Fesenmaier, D., Klein, S., Buhalis, D. (eds.) Informaton and Communication Technologies in Tourism 2000: ENTER 2000, pp. 15–27. Springer, Wien (2000)

Mann, W., Thompson, S.: Rhetorical structure theory: a theory of text organization. T. R. ISI/RS87–190, USC/ISI, Marina del Rey, CA (1987)

McCrae, R.R., John, O.P.: An introduction to the five-factors model and its applications. J. Personality **60**, 175–215 (1992)

McKeown, K.R.: Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text. Cambridge University Press, London, UK (1985)

Nardon, M., Pianesi, F., Zancanaro, M.: Interactive documentaries: first usability studies. The 2nd Workshop on Personalization in Future TV, Malaga, Spain, available online at: http://itv.eltrun.aueb.gr/topics/ah02tv02/ (2002)

Niederee, C., Stewart, A., Mehta, B., Hemmje, M.: A multi-dimensional, unified user model for cross-system personalization. Workshop On Environments For Personalized Information Access, pp. 34–54. Gallipoli, Italy (2004)

Norman, D.A.: The Invisible Computer. MIT Press, Cambridge, MA (1998)

Not, E., Petrelli, D., Sarini, M., Stock, O., Strapparava, C., Zancanaro, M.: Hypernavigation in the physical space: adapting presentations to the user and to the situational context. New Rev. Hypermedia Multimedia **4**, 33–45 (1998)

Novak, T.P., Hoffman, D.L., Yung, Y.F.: Measuring the customer experience in on-line environments: a structural modeling approach. Marketing Science **19**, 22–44 (2000)

Oberlander, J., O'Donnell, M., Knott, A., Mellish, C.: Conversation in the museum: experiments in dynamic hypermedia with the intelligent labelling explorer. New Rev. Hypermedia Multimedia. **4**, 11–32 (1998)

Petrelli, D., Not, E.: User-centered design of flexible hypermedia for a mobile guide: reflections on the hyperaudio experience. User Model. User-Adapt. Interact. **15**(3–4), 303–338 (2005)

Popovich, P.M., Hyde, K.R., Zakrajsek, T., Blumer, C.: The development of the attitudes toward computer usage scale. Educ. Psychol. Meas. **47**, 261–269 (1987)

Potonniee, O.: Ubiquitous personalization: a smart card based approach. The 4th Gemplus Developer Conference, Singapore, available online at: http://www.gemplus.com/smart/rd/publications/pdf/Pot02per.pdf. (2002)

Raptis, D., Tselios, N., Avouris, N.: Context-based design of mobile applications for museums: a survey of existing practices. The 7th International Conference on Human Computer Interaction with Mobile Devices & Services, pp. 153–160. Salzburg, Austria (2005)

Reiter, E.: Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible? The Seventh International Workshop on Natural Language Generation, pp. 163–170. Kennebunkport, ME (1994)

Rocchi, C., Zancanaro, M.: Adaptive Video Documentaries. The HyperText Conference, pp. 36–37. Nottingham, UK (2003).

Rocchi, C., Zancanaro, M.: Template-based adaptive video documentaries. Artifical Intelligence in Mobile Systems, pp. 79–83. Nottingham, UK (2004)

Rocchi, C., Stock, O., Zancanaro, M., Kruppa, M., Krüger, A.: The museum visit: generating seamless personalized presentations on multiple devices. The 2004 Conference on Intelligent User Interfaces, pp. 316–318. Madeira, Portugal (2004)

Rotter, J.B.: Generalized Expectancies for Internal Versus External Control of Reinforcement. Psychol. Monogr. **80**, 1–28 (1966)

Sarini, M., Strapparava, C.: Building a user model for a museum exploration and information-providing adaptive system. In: Brusilovsky, P., De Bra, P. (eds.) The 2nd Workshop on Adaptive Hypertext and Hypermedia, Report No. 98/12, CS Reports, Eindhoven University of Technology, Eindhoven, The Netherlands, pp. 63–68 (1998)

Shneiderman, B.: Eight Golden Rules for Interface Design. Designing the User Interface, 3rd edn. Addison Wesley, Reading MA (1998)

Stock, O.: The ALFRESCO Project Team.: ALFRESCO: enjoying the combination of NLP and hypermedia for information exploration. In: Maybury, M.T. (ed.) Intelligent Multimedia Interfaces, pp. 197–224. AAAI Press, Menlo Park, CA (1993)

Stokmans, M.J.W.: Reading attitude and its effect on leisure time reading. Poetics **26**, 245–261 (1999)

Totterdell, P., Boyle, E.: The evaluation of adaptive systems. In: Browne D., Totterdell P., Norman M. (eds.) Adaptive User Interfaces, pp. 161–194. Academic Press, London (1990)

van Mulken, S., André, E., Mueller, J.: The persona effect: how substantial is it. Human Computer Interaction Conference, pp. 53–58. Berlin, Germany (1998)

Want, R., Schilit, B., Adams, N., Gold, R., Petersen, K., Ellis, J., Goldberg, D., Weiser, M.: The PARCTab ubiquitous computing experiment. T.R.CSL-95–1. Xerox PARC, Palo Alto, CA (1995)

Authors' vitae

Oliviero Stock has been at IRST since 1988 and has been its Director from 1997 to 2001. His work is mainly in artificial intelligence, natural language processing, intelligent user interfaces, and cognitive technologies. He is the author of 180 papers, author or editor of twelve volumes, and a member of the editorial board of a dozen scientific journals. He is a fellow of ECCAI and AAAI. He has worked on intelligent interfaces for cultural heritage appreciation for nearly two decades, has been the coordinator of PEACH and currently co-ordinates the IRST-University of Haifa Collaboration Project.

Massimo Zancanaro has been working as a researcher in the Cognitive and Communication Technologies Division at ITC-irst since 1994. He received is Laurea Degree in Computer Science from the University of Milano in 1992. His primary interest is in the field of Intelligent Interfaces and in particular language-based interfaces. Since 2004 he has taught Computer-Human Interaction at the International PhD in ICT program at the University of Trento.

Paolo Busetta is currently a software consultant in the area of service oriented systems. Until recently, his research focused on architectures and languages for multi-agent systems, in particular in the areas of teamwork and resilient systems. He has a Master degree in Computer Science from the University of Melbourne. He has a long professional experience in software development accumulated in over 26 years of activity.

Charles Callaway is currently a research scientist in Informatics at the University of Edinburgh working on applications of natural language technology. He received his Bachelor degrees in Computer Science, Mathematics and Latin along with a Masters in Computer Science from the University of Texas at Austin, and completed his Ph.D. in 2000 at North Carolina State University. Dr. Callaway has worked in several areas of AI and Computational Linguistics including natural language generation, embodied agents, tutoring systems, story generation, knowledge representation and language-based information presentation.

Antonio Krüger is an associate professor at the Institute for Geoinformatics and Computer Science of the University of Münster. He received his diploma and doctoral degree in Artificial Intelligence from Saarland University, Germany. Dr. Krüger is working in the area of intelligent user interfaces, user modeling, smart environments and mobile information systems and is especially interested in the influence of ubiquitous computing on these fields. He is also one of the co-founders of Eyeled Ltd, a company dedicated to the design of adaptive mobile and ubiquitous information systems.

Michael Kruppa is a researcher at the German Research Center for Artificial Intelligence (DFKI). He received his diploma and doctoral degree in Artificial Intelligence from Saarland University, Germany. Dr. Kruppa is working in the area of automatic usability evaluation, intelligent user interfaces, virtual characters and mobile computing. He is currently working for the MeMo project funded by the German Telekom and also in the Compass 2008 project for the Olympic Games 2008.

Tsvi Kuflik is a Lecturer in the Management Information Systems Department at the University of Haifa. He received his BSc. and Msc. in computer science and Ph.D. in Information Systems from the Ben-Gurion University of the Negev, Beer Sheva, Israel. Before joining the University of Haifa he worked as a researcher at IRST in Trento, working on user modeling in active museums, work he continues at the University of Haifa in the framework of the collaboration of the University of Haifa and IRST where he leads a research group focusing on supporting museum visitors. Dr. Kuflik research interests focus mainly on user modeling, ubiquitous user modeling and information filtering.

Elena Not received her degree in Computer Science from the University of Udine (Italy). Since 1992 she has been a Research Scientist at the Cognitive and Communication Technologies Division at ITC-irst. She worked

on several national and international research projects reflecting her interest in natural language generation with particular focus on text planning, discourse structure, referring expressions generation and disambiguation, dynamic hypermedia generation, and integration with other output modalities. The joint research described in this volume reflects her substantial experience in the development of adaptive multimedia information presentation systems.

Cesare Rocchi is currently a research consultant with IRST. He received a degree in Philosophy from the University of Bologna in 2001. He has been working in the areas of Natural Language Generation, Adaptive Hypermedia and User Modeling. He is also a PhD student at the artificial intelligence department of the Università Politecnica delle Marche, working on the design and evaluation of human-machine interfaces.