Multimodal Audio Guide for Museums and Exhibitions

Sandra Gebbensleben, Jana Dittmann, Claus Vielhauer
Otto-von-Guericke University of Magdeburg, Institute of Technical and Business Information Systems, Research Group on Multimedia and Security, Magdeburg 39106, Germany

ABSTRACT

In our paper we introduce a new Audio Guide concept for exploring buildings, realms and exhibitions. Actual proposed solutions work in most cases with pre-defined devices, which users have to buy or borrow. These systems often go along with complex technical installations and require a great degree of user training for device handling. Furthermore, the activation of audio commentary related to the exhibition objects is typically based on additional components like infrared, radio frequency or GPS technology. Beside the necessity of installation of specific devices for user location, these approaches often only support automatic activation with no or limited user interaction. Therefore, elaboration of alternative concepts appears worthwhile.
Motivated by these aspects, we introduce a new concept based on usage of the visitor's own mobile smart phone. The advantages in our approach are twofold: firstly the Audio Guide can be used in various places without any purchase and extensive installation of additional components in or around the exhibition object. Secondly, the visitors can experience the exhibition on individual tours only by uploading the Audio Guide at a single point of entry, the Audio Guide Service Counter, and keeping it on her or his personal device. Furthermore, since the user usually is quite familiar with the interface of her or his phone and can thus interact with the application device easily. Our technical concept makes use of two general ideas for location detection and activation. Firstly, we suggest an enhanced interactive number based activation by exploiting the visual capabilities of modern smart phones and secondly we outline an active digital audio watermarking approach, where information about objects are transmitted via an analog audio channel.

Keywords: Multimedia Audio Guide, Human-Computer-Interaction, Mobile Smart Phone, Smart Phone, Usability, Digital Audio Watermarking

1. INTRODUCTION

This section gives an introduction to our new approach for a Multimodal Audio Guide and outlines major problems of the already existing systems by introduction of design criteria for a Multimodal Audio Guide and comparison of these with existing systems.
The typical Audio Guide application, which is used in today's museums works as follows: The visitor usually borrows an electronical device at the Audio Guide Service Counter of the museum. On this device, different audio commentaries are stored, which are related to special exhibits. If the visitor wants to access more information about these special exhibits, she or he starts the device and the audio commentaries are played. These devices usually have different interface design and different activation techniques. The activation can be divided in two main groups: *manual activation* and *automatic activation*. In the *Manual activation* case the visitor herself or himself initiates playback of the commentary directly by entering an object identification number, for example. For the *automatic activation*, typically different wireless technologies such as infrared, radio frequency or GPS (General position System) are used. Here, a signal is permanently broadcasted to the receiver of the Audio Guide device and if the visitor enters the vicinity of the related object than the audio commentary is started automatically in accordance to the received signal. Often these devices only have a small display, where only a very limited number of characters or digits regarding the information can be shown, for example object identification numbers or titles of the related object.
From this analysis we conclude, that we observe in general the following disadvantages:

1. provision and distribution of the Audio Guide devices,
2. installations of activation sensors,
3. training of device handling,

Multimedia on Mobile Devices II, edited by Reiner Creutzburg, Jarmo H. Takala, Chang Wen Chen,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6074, 60740S, © 2006 SPIE-IS&T · 0277-786X/06/$15

4. interface design not always conform with human-computer-interaction (HCI) and
5. no or limited visual feedback during usage of the Audio Guide.

Therefore, in our research, we investigate different approaches of a new design and prototypically implement a new Audio Guide concept based on the mobile phones of the visitors, where we mainly look into mobile smart phones due to the overall required hardware and software functionality. The one hand, a main advantage of our approach is, that the visitors are familiar with their devices and do not have to practice to use an entirely new device. On the other hand, this solution saves money for the museum, because they do not have to purchase Audio Guide devices for lending to their visitors. To be more flexible in starting audio comments during the tour, we suggest a combined manual and automatic activation method.

In our new design, we are furthermore motivated by an application domain, where visitors use the Audio Guide inside the building as well as in the outside surroundings, like is the case for example for the *"Meisterhäuser"* in Dessau [1]. This realm consists of three conserved buildings, which are recognized as UNESCO world cultural heritage. The buildings show interesting architecture and design in the interior and the outer area, and consequently, because of the cultural heritage, the buildings are under monumental protection. Therefore only very limited modifications of the buildings and technical installations are allowed. Another limitation with these buildings is that the doors are too small for access by persons in wheel chairs. This means that impaired people sitting in a wheel chair have no possibility to physically access the interior of the buildings. Based on this application scenario, we have identified the following technical and usage criteria for Audio Guides: The following **technical criteria** concentrate primarily on the interests of the museum, limitations by monument conservation and technical installations:

1. no purchase, provision and distribution of devices,
2. no installations of localization sensors,
3. support for both interior and outer area of the museum,
4. update functionality of information and
5. no additional devices, Audio Guide software shall work on widespread personal devices.

Audio Guides with these properties, offer a cheap and easy solution for each museum, because they do not need to buy and install anything additionally. The staff at the Audio Guide Service Counter can control the distribution of the Multimodal Audio Guide software .
The following **usage criteria** focus mainly on the interests of the visitor. Their intension is to make the usage of the Audio Guide easy and thus the visit at the museum more enjoyable. The Audio Guide should fulfill the following criteria :

1. multilingual tours,
2. different tours for different groups (e.g. Parents and Children),
3. individual tours,
4. easy handling of the device and
5. multimedia support of content.

By fulfilling the above properties, the Audio Guide is adequate for visitors interested in a *guided tour*. The *multilingual tours* allow presenting the content in different languages and focusing on international visitors. It appears reasonable to create *different contents* for *different groups of visitors*, because for example children usually do not understand the information if it is given to them like to the adults. It should furthermore be also possible to create content with different degrees of granularity for *disabled people* like visually impaired people or people sitting in a wheel chair, who could not always reach each point in the museum or exhibition. The easy handling, or *usability*, of a device is a basic requirement of each electronically device and it helps also older people to use the Multimodal Audio Guide. Further, if the Audio Guide supports the audio comments with *textual representation* and *pictures or animations*, this helps for navigation through the exhibition and for presenting the content more closely to the visitor.

In the following, the problems with the existing Audio Guide applications and the concept of Audio Guides in general are discussed to further motivate our new approach. In our discussion, we look at four main aspects: complexity of technical installations, user training, activation and HCI, and discuss the capability of the existing application with respect to these aspects.

The *complex technical installations*, which are necessary for most automatically activated devices of the current Audio Guides, raise problems in museums and exhibitions, if these are located in cultural heritage buildings like the *"Meisterhäuser"* in Dessau [1]. Here, due to reasons such as monumental conservation, it is often given that it is not allowed to perform any modifications in the structure and furnishing of these buildings. Therefore, very often, in these cases Audio Guides are not or only limited offered due to technical constraints. However, usually technical installations are acceptable for the *distribution* of the Audio Guides and for the usage in the exhibition in the reception area. This allows setting up a Audio Guide Service Counter to maintain and to distribute the Audio Guides. For the adjustment of a special tour and language some of the used devices need an additional device at this Audio Guide Service Counter, like an infrared hotspot. For the activation of the Audio Guide during the tour additional devices become necessary. This is required mainly for the automatic activation of the audio comment. Other kinds of Audio Guides use a combination of radio frequency and wireless local area networks. For these systems, also installations of location devices are necessary.

The second aspect, which we elaborate in this paper, is the *usability*, which can be divided in user training for the activation and user training for the usage of the Audio Guide. The visitors borrow an Audio Guide device at the Audio Guide Service Counter of the museum. Often they are exposed to devices, which they have never used before. Therefore they have to get acquainted to how to use the Audio Guide. It is not always obvious for each person how to activate the Audio Guide and how to use it for activating the audio comments. Many of the existing Audio Guides have no or only a limited display to assist the people using the Audio Guide, for example telling them with animations which button to press for activation.

For *activation* of the audio comments, three different ways are possible: manual, automatic, manual and automatic activation. If only manual activation is chosen for the Audio Guide, the visitor always has to become active to initiate the audio comments, in most cases by entering an appropriate object number. This can be very exhausting for the visitors, especially if the object numbers are not fixed at the objects. Often in these scenarios the visitors have than to carry a paper document like a catalogue or map in order to identify the related number to an exhibit. For automatic activation, different techniques are possible and the audio comments are started if the visitor enters a predefined area near the related object. This is why in these cases the visitor can only access the actual comment related to the object close to her or him. However, some visitors might like to hear again an earlier commentary from a previous exhibit. Hence the visitor is required to go back to the previous object, which is not very comfortable. Therefore a combined activation seems to be the best solution. In this case, the visitor is automatically triggered and additionally has the control on which audio track she or he can start, how often and at which location of the exhibition it shall be replayed. This also allows the visitors to plan their routes through the exhibition, because they can listen ahead to the comments from exhibits they have not visited yet.

The last mentioned aspect is the *human-computer-interaction*. Often, the existing Audio Guides have a display rather limited in size, where only a few digits or words are visible. This makes it difficult for the visitors to see if they have performed some faulty operation, e.g. pushed a wrong button or where the fault is, in case the activation of an audio comment does not function. Furthermore, they do not get any supporting function to plan their route or to find a special part of the exhibition or related objects in the exhibition. By supporting a display, which can also show pictures and more amount of text, it is possible to provide position information to the visitor, content such as a preview of what she or he can see in the next rooms, error conditions and explanations why a comment does not start and many more. Consequently, for a good Audio Guide design, an adequate display appears essential.

2. CURRENT AUDIO GUIDES IN COMPARISON

In this section of our paper, we firstly summarize the currently most used techniques, which are identified in our work, as well as their device properties. Subsequently, we compare the properties of the Audio Guides with the criteria identified in the previous section of the paper.
The Audio Guides, which are often used in museums or exhibition, can be categorized into the subsequent categories:

1. devices with number pads,
2. personal Digital Assistants (PDA),
3. mobile phones and
4. devices for special scenarios.

The devices equipped with number pads have a similar appearance as regular wireless telephones. They differ in size and design, but have similar functions. These devices are in most cases manually activated. The scenario is as follows: The visitor goes trough the exhibition or the museum and starts the audio commentary by typing in the appropriate object number. Some of these devices can also be activated automatically by using technologies like infrared and radio frequency. Three examples of this group are from PRO Cept GmbH the Mediaexplorer [2], the Museum Exhibit Guide [3] and from Antenna Audio the X-Plorer [4].
The second group of these Audio Guides uses the widespread PDA devices. Here two possible activation methods exist: manual or automatic activation. The object numbers can be entered by using the touch screen and the automatic activation can use different wireless technologies. With these devices additional information to the audio commentary, can be presented using visual representation on the PDA display, too. Some of the providers already use this possibility, but not all. Today, none of these PDA-based Audio Guides are available for use by the visitors on their private PDAs. Rather than that, the visitors have to borrow the PDA at the reception desk. Some examples of this category are the PDA based Guides from eloqu – metabasis GmbH [5] and from cool IT GmbH the coolMuseum [6].

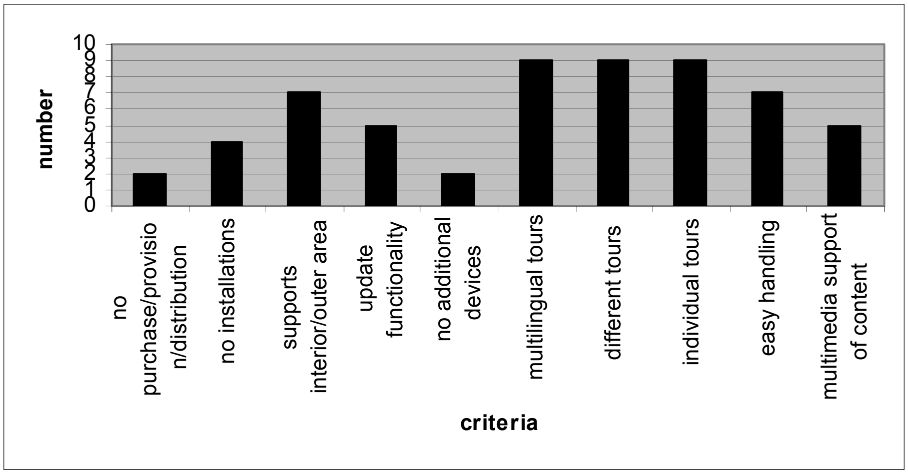
A small group of providers exist, who are already addressing mobile phones to present the visitor more information about special exhibits or buildings. Known to us are at the moment BeyondGuide [7], Spatial Adventures [8] and Touch Graphics [9]. Until now, there appears only one single activation method for this category: manual activation. The visitor calls a special telephone number and gets connected to an automated telephony application service. Here, all audio commentaries and information are centrally stored and accessible for the user via an interactive telephony-based user interface. The audio content is transmitted via the speech channel of the telephone connection. Examples for this category of Audio Guides are BeyondGuide [7] and Rocky Mountains Audio Tours [10]. A disadvantage of these Audio Guides from the perspective of visitors is the fact that they typically have to bear the charges for the call additionally to the Audio Guide fee. However, on the other side this method may provide a simple business model for the service provider.

The last category of devices, which are mostly designed for special scenarios, uses specifically designed mobile devices to present the information to the visitor. These devices differ in look and functionality, also the concept of activation is different. Some can be activated automatically, manually or automatically and manually. Examples of this group are the Sennheiser Guide Port [11] and the dataton PickUp version 1.2 [12]. The Sennheiser Guide Port uses radio frequency and a wireless local area network (WLAN) for the automatic activation of the audio commentaries. The installation of identifiers is necessary for each exhibit. The dataton PickUp version 1.2 uses infrared hotspots for the activation of audio comments. Here the visitor points with the Audio Guide device on a hotspot near the object and pushes the play button to activate the related audio information.

In the following, we perform a comparison of the mentioned Audio Guides with respect to our criteria from section 2. Many Audio Guides implementations exist and it is not possible to mention all of them in this paper. In our limited study, we thus compared 47 selected Audio Guides from different providers based on a web inquiry. In this paper nine of these Audio Guides are mentioned, at least two for each group. The devices with number pads only differ in size and design, but they have almost identical functionality. This is the reason why only three examples from this category are selected to compare these with our criteria. This is also true for the other categories. Mobile phones and PDAs have all very similar possibilities to present information to the visitors and we therefore select two of each category as representative approaches. For the category of special designed devices, two are pick out from all evaluated Audio Guides, which represent the typical properties of all Audio Guides from this category.
Figure 1 summarizes the result of the comparison of existing Audio Guides. Only three criteria are fulfilled by all nine Audio Guides: multilingual, different tours for children, impaired people and individual tours. Two Audio Guides work without distribution of devices and additional devices. These two are from the category, which already uses the visitor's mobile phone to make more information available to the visitors. Two more devices work without any physical installations in the museum or exhibition. These devices are activated manually and if no further installations are present, the visitor has to carry along a list with the object numbers. More than the half of the Guides could be easily

handled, and these devices are mainly from the groups “PDA” and “mobile phone”, with an additional multimedia support. From the remaining Audio Guides, only a few have a full display to support the audio information with pictures and textual representation.

Figure 1: Comparison of existing Audio Guides



The results of our comparison of the existing Audio Guides with our previously defined criteria show that today no Audio Guide can be identified, which fulfills all of the defined properties. This is true for all 47 evaluated Audio Guides. Consequently, we decided to develop new Audio Guide hardware concept based on a software application to be executed on the visitor’s mobile phone, with the goal to make the concept attractive for the museums as well as for visitors. It should be attractive for museums and exhibitions, since they are not required to purchase and maintain any Audio Guide hardware and they don’t need to install localization devices. For the visitors our concept appears attractive, because they can use their personal mobile phone, a device they should be able to handle very well and in case of smart phones, can support the audio information with additional media representations of text and image.

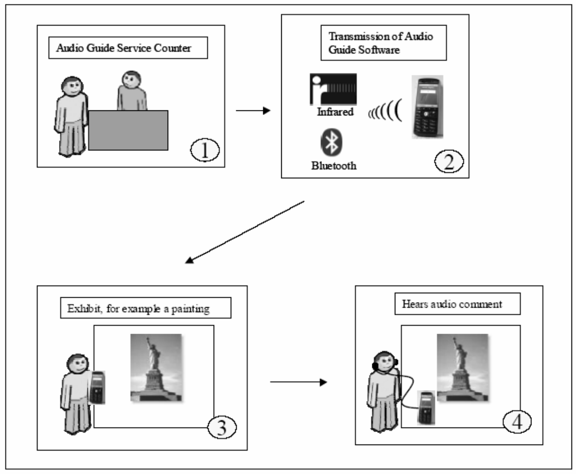
3. DESIGN AND CONCEPT OF THE MULTIMODAL AUDIO GUIDE

The basic idea of our multimodal Audio Guide is based on the criteria introduced in section 1 to fulfill all technical and usage criteria by introducing a guide for mobile smart phones with a combined manual and automatical activation of audio comments. Firstly the overall scenario can be described as follows:

The visitor uses his personal mobile smart phone, she or he has to go to the Audio Guide Service Counter of the museum or exhibition and gets the Audio Guide software transferred to his mobile smart phone. The transformation can be performed via wireless interfaces such as Infrared or Bluetooth. After this transfer, the visitor can install the software on her or his mobile smart phone, start the application and begin the tour through the exhibition. At specified points, which can be rooms, outdoor locations or exhibits, the visitor can either activate the Audio Guide manually by entering the object number, or the automatic activation starts the audio information. The Audio Guide works generally inside and outside of buildings. The update functionality is implicitly integrated, since each visitor gets the up to date version of the Audio Guide software transferred to the mobile smart phone at the time of software transfer. For the museum a management platform is offered to update the information.

Figure 2 below illustrates our Audio Guide scenario. At the position identified with the yellow number 1, the visitor enters the museum at the Audio Guide Service Counter. The Audio Guide software gets transferred to the mobile smart phone via one of the wireless interfaces, Infrared or Bluetooth (number 2). Then the visitor starts the individual tour trough the exhibition. In case she or he wants to get additional information about an object, the audio comments can be started by entering the appropriate object number or the audio comments start automatically (number 3). The visitor can hear the audio comments by either using headphones connected to the mobile smart phone or using the integrated mobile smart phone speaker (number 4).

Figure 2: Concept of the Audio Guide



Secondly as already discussed the new Audio Guide software can be activated manually and automatically. We combined both possibilities of activation, because we want to combine the advantages of both. In areas, where automatic activation is not possible or not desired, the visitor types in the related object number.

For the manual activation the visitor has to enter the object number, which can be seen from graphical and textual information on the mobile display, e.g. a floor plan for navigation. Thus the Audio Guide supports multimedia presentations.

For the automatic activation we suggest to make use of digital watermarking techniques in sound systems, as will be explained in more detail in section 3.3. For these activation methods, no additional devices like localization devices are required. In the software design, we have considered aspects of human-computer-interaction in such way, that the tours through the exhibition have different content for different visitors and that the visitor can create his individual tour.

In the following subsections we outline the system requirements and introduce the two different activation approaches of our Audio Guide software.

3.1. System requirements

In order to fulfill the usage criteria from section 1, our new Audio Guide software is required to have the properties listed below. These properties are related to the overall capability of mobile smart phones to load and execute additional user-specific software. Our design requires the following device properties:

- infrared and/or Bluetooth interface,
- relatively large storage capacity,
- software environment (application programming interface, API) for Audio Guide application,
- integrated Audio Player,
- speaker or headphones and
- microphone

For deployment one of the *interfaces***,** Infrared or Bluetooth are required to transfer the Audio Guide software from the Audio Guide Service Counter to the visitor’s mobile smart phone. By using wireless interfaces, it is not required to provide standardized cable connections, which may vary widely from one type of mobile smart phone to another. Today, many mobile smart phones offer one or both of the mentioned wireless interfaces.

The requirement regarding *storage* capacity depends on how many points are defined where information should be presented and how much information should be presented to the visitor. For example, a small exhibition where only audio and textual information should be presented will need less memory than a large exhibition where additionally to audio and textual information also pictures and maps should be presented. In our scenario of the “Meisterhäuser” in Dessau, we have for nine objects with audio commentaries in AMR audio format, maps and textual representations about 400 Kilobytes.

For installation and deployment of the Audio Guide *software* a runtime environment is required. A standard software environment for mobile phones is J2ME (Java Two Mobile Edition). The mobile smart phones also need an integrated *Audio Player* to play the audio comments. One common feature of many mobile smart phones today is an integrated MP3-Player, however other audio formats shall also be considered. Furthermore and quite obviously, all smart phones are equipped with a *speaker* and a *microphone*; however for our approach we require that these devices are accessible by application programs via appropriate device interface drivers in the phone's operating system.
All mobile smart phones, which fulfill the above criteria, can be used as Audio Guide. The problem we face in practice is the fact that there exists Java APIs, which support MP3 and other audio formats for the audio representation, but it apparently, not all providers have implemented these APIs between the Java environment and the Audio Player. Therefore the Audio Player from the mobile smart phone is not always capable to replay the audio comments from the Audio Guide software from the same audio file format. Here it is a technical challenge to us, to extend our software in this way that the audio replay function will work adaptively on different mobile smart phone, by automatic adjustment to the corresponding file type. The table below summarizes the compatibility list of mobile smart phones from our first evaluation (test), which can be used with our application, including audio support for playing the comments. However, we have not attempted to evaluate all mobile smart phones on the market today with respect to compatibility to our software, because of the large quantity of different types and the great dynamics in the technical development in this area.

Table 1: List of mobile smart phones

Brand	**Model**	**MIDP Version**	**Successful**
Motorola	V3	2.0	✓
Nokia	6230	2.0	✓
Siemens	SL65	2.0	✓
Sony Ericsson	T630	1.0	✓
Sony Ericsson	K700i	2.0	✓

3.2. Design of manual activation

The concept of the manual activation is well-known and is used also for many existing Audio Guides. For the manual activation of the Audio Guide, the visitor has to key in the appropriate object number. The object number points to a specific audio comment, which has been transferred to the mobile smart phone along with software. Therefore the audio comment is clearly related to a specific exhibit and can be started directly.
The manual activation concept together with the idea of using the visitor's mobile smart phone fulfills all our defined technical and usage criteria. It combines two different ideas: the number based activation on the visitor's private mobile smart phone. The new parts of this approach are firstly to store the audio and multimedia representations of the information locally on the mobile smart phone and not on a centralized server and secondly to use the display of the mobile smart phones to support the object identification number information with visual representations. Figure 3 below shows one possible prototype display to present the visual information for human-computer-interaction. On the top area of the display, the visitor can see a map with his marked position and below the map the textual representation. By pushing the buttons left or right, the user can navigate back to the entry screen or re-start the audio comments manually.

Figure 3: prototype display of the Audio Guide software

This manual activation is a long time practiced and proven technique, but not very comfortable to the user, because it is not always possible to fix the activation numbers directly near the object. As already mentioned for some Audio Guides, the visitor needs to carry along a list with all exhibits and the related numbers. To make the usage of the new Audio Guide more comfortable to the visitors, we decided to combine the manual activation with an automatic activation method and an extended visual feedback regarding position and operation of the Audio Guide.

Figure 4: Graphical scheme of manual activation

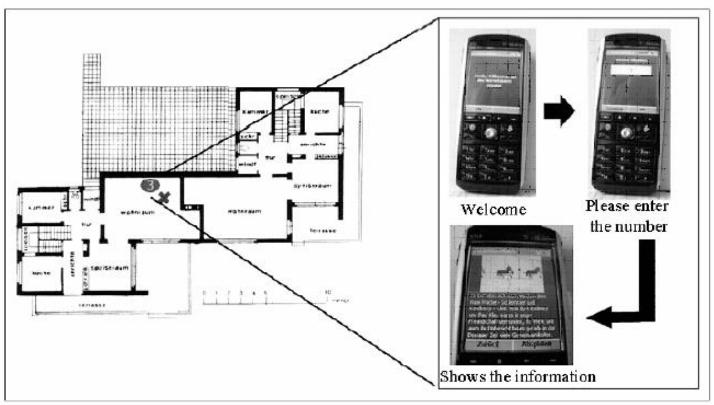


Figure 4 outlines the graphical scheme of our manual activation approach. If the visitor starts the Audio Guide software the first time, she or he is welcomed to the exhibition and the software displays the welcome screen (left top image of mobile smart phone: Welcome). As the visitor walks through the exhibition she or he may find an exhibit to know more about. In our example, it has the object identification number '3', which is shown in the map on the left side of figure 4. The user then types in the object number and presses the start button (right top image of mobile smart phone: Please enter the number). Now, a map is displayed on the mobile smart phone display, with information about the location in the exhibition and the textual representation of the information. By pushing the 'play'- button the visitor may listen to the audio commentary, as shown in the lower image (lower image of the mobile smart phone: Shows the information).

3.3. Watermarking based activation

For the automatic activation we decided that it is not recommendable to use one of the techniques, which are already used for the existing Audio Guides, because these techniques often go along with additional devices and complex installations. Our approach for automatic activation is therefore to use digital audio watermarking [13], making use of the condition that museums often already have loudspeakers in the exhibition rooms for presenting sound information or to make announcements. Digital watermarking techniques based on steganographic systems offer the possibility to embed information directly into the media data. Most watermarking techniques are known from the copyright

identification, owner or customer authentication or integrity and manipulation recognition, see for example in [14] or [13]. In our case we use digital watermarking as annotation or illustration watermarks, known from the literature see for example in [15].
Our scenario is as follows: In specified rooms or near specific exhibits a soft sound is played. In this sound a digital audio watermark is embedded. The idea is to use the watermark as object identification number (option 1: object identification watermark *OID*). Additionally with a high watermark capacity, the watermark can be used to transmit the object information, which should be presented to the visitor (option 2: object information watermark *OIW*). The visitor's mobile smart phone recognizes the played music through the integrated microphone and analyses the sound. The watermark is detected by a watermarking detection algorithm on the mobile smart phone (part of the transferred software) and the related information is started.

Figure 5: Graphical scheme of the watermark activation

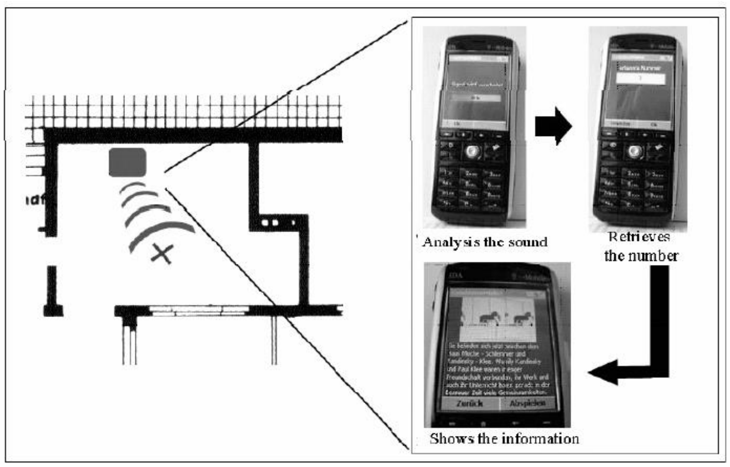


Figure 5 above illustrates the activation based on digital watermarked *OID* sound. The visitor comes in a room where sound is played. The mobile smart phone records the sound with the microphone and starts to analyze the embedded watermark to find the *OID* (top left image of the mobile smart phone: Analysis the sound). The watermark in this case contains only the identification object number. The mobile smart phone shows in our example the number '3' (top right image of the mobile smart phone: Retrieves the number). The screen with the map and the textual representation becomes visible (lower image of the mobile smart phone: Shows the information) and the visitor has now the same possibilities as with the manual activation; read the text and/or hear the audio information.

By using digital audio watermarking we have to consider the following parameters known from digital watermarking algorithms: robustness and fragility, security, transparency, complexity, capacity, verification and invertibility

For our application scenario we therefore need to determine the parameter requirements and select an appropriate watermarking algorithm for the required parameter settings from variety of existing techniques, see for example introduced algorithms from watermarking benchmarking in [12].
The parameter **robustness and fragility** determines if the watermark can be detected after media operations and malicious or non-malicious attacks. The fragility determines in which cases the watermark can be destroyed or removed successfully. In the applications for copyright, owner or customer identification we usually require a robust watermark. For the manipulation recognition and integrity checks we usually need a fragile watermark. In our case of annotation watermarking we need a robust watermarking approach to be resistant to digital-analog and analog-digital conversions and additional noise, which is recorded together with the watermarked sound at the mobile smart phone. The digital sound is watermarked in advance by embedding the *OID* and optionally the *OIW* and played in the exhibition. The mobile smart phone records the analogue sound in front of the object and analyses the recordings in the digital domain to find *OID* and optionally the *OIW* watermark.

Security describes whether the embedded watermarking information cannot be removed beyond reliable detection by targeted attacks based on a full knowledge of the embedding algorithm and the detector, except the key, and the knowledge of at least one watermarked data. In our scenario we do not consider security aspects as the played sounds are only used as information carrier. If the sound itself needs to be protected an additional copyright watermark is required.

The **transparency** relates to the properties of the human sensory. A transparent watermark causes no artifacts or quality loss due to embedding the watermark. In our case it means, that no artifacts are heard in the played sound. This is important, because it would disturb the atmosphere in the exhibition and the visitor could be annoyed by pure audio quality.

Complexity describes the time and effort, which is needed to embed and retrieve a watermarking. For us only the time for retrieval is essential, because the embedding can be performed in advance where we do not have real-time constraints. The watermark retrieval on the mobile smart phone should be performed closed to real time with minimal delay, because the visitor shall get the information in the moment he stands in front of the exhibit.

The **capacity** of a watermarking describes how much information can be embedded, how often the watermarking can be embedded or how many different watermarks can be embedded in parallel in the audio file. This is a very important parameter for our application, which sometimes also called payload. The watermark capacity has to be at least as high as to embed the *OID* (which needs log_2*(#number of objects)*) bits). For example, if we have 200 objects in the exhibition we need at minimum 8 bits and for 3000 at minimum 12 bits. Here two main different kinds of watermarking techniques are of interest: A) The **watermarking pattern approaches**, by embedding the *OID* as unique watermark pattern representing the object identification number (during retrieval we would search for the pattern and receive a binary decision if the pattern was found). B) The **codeword watermarking approach** by embedding the *OID* directly as for example binary number watermark information. Codeword watermarking approaches allow the embedding of binary strings and if the capacity is large enough also the embedding of further *OIW* information like text, audio commentary and pictures. The watermark channel in this later case is used to transmit the annotations directly.
Furthermore to support different tours the *OID* can be enhanced with different trigger flags and allows us to embed different information in parallel to activate the Audio Guide appropriately. For example the mobile smart phones of the different user groups like adults and children can be activated by different OID trigger flags using the same sound. Additionally copyright information can also be embedded if the capacity is large enough.

The **verification** procedure describes the key constellation: private verification like private key functions or a public verification possibility like the public key algorithms in cryptography. For our scenario verification with one pre-initialized key setting is appropriate, as we do not consider security aspects. **Invertibility** describes the possibility to produce the original data during the watermark retrieval, which is not required in our application.

From the overall discussion we see, that the most important parameters are: robustness against D/A and A/D conversions as well as adding background noise, transparency to the human ear, low complexity during retrieval, capacity for *OID*, for *OID* with additional trigger flags and in accordance capacity for *OIW*.

From the watermark design we know the optimization of the parameters is mutually competitive and cannot be clearly done at the same time. If on the one hand we want to embed a large message, we cannot require large robustness simultaneously. A reasonable compromise is always a necessity. On the other hand, if robustness to strong distortion is an issue, the message that can be reliably hidden must not be too long. Benchmarking results in [16] show for five different audio watermarking algorithms a comparison of complexity, robustness, transparency and capacity. The test results from this benchmarking show, that the five evaluated watermarking algorithms have different properties. Depending on the working domain, the evaluated test results differ for all tested parameters. Furthermore, the dependency on the properties of digital watermarks is proved. From the general perspective if an algorithm provides a high robustness, then the transparency decreases (like for VAWW – Viper Audio Water Wavelet, which is a watermarking algorithm working in wavelet domain by embedding the watermark in selected coefficients, see [16]). For example the provided transformations DynamicPitchScale, FFT_Stat1, Pitchscale and Resampling, see in [16], disable or even remove the watermark successfully for all evaluated watermarking algorithms and all tested audio files (total number of 389 files, used watermark message: *UniversityOfMagdeburg*). This shows that an appropriate watermarking algorithms and parameter selection needs to be determined on further tests and we expect that the existing approaches needs to be tuned and improved for our application scenario.

4. CONLUSIONS AND FUTURE WORK

We have explained the general concept of electronical Audio Guides and which advantages and disadvantages are related. To address the main weaknesses in existing Audio Guide approaches, we introduce new technical and usage criteria for Audio Guides, based on an exemplary implementation object, the "*Meisterhäuser*" in Dessau, Germany [1]. Details about existing Audio Guides and their properties are discussed and compared to our technical and usage criteria. The result of the comparison is, that no existing design for an Audio Guide exists, which can fulfill all our criteria. Therefore we decided to develop a new Audio Guide system concept and identify the system requirements to perform well. In our approach, we deploy the new software on a private mobile smart phones based on smart phone technology, which to an increasing extent posses properties, that are sufficient for executing the Audio Guide software. Regarding the activation method, we introduce a combined approach of manual and automatic activation. The manual activation only requires the number pad on the mobile smart phone, whereas for the automatic activation, we utilize the microphone of the device, to record and analyze watermarked sound from the environment of the exhibition object. The main technical and implementation challenge we have here are the application interfaces between the Java environment on the mobile smart phone and the integrated Audio player. From the first test we see, that unfortunately today, not too many mobile smart phones implement this interface.

With our new Audio Guide approach a new and innovative system is given to museum and exhibitions, which can be used with less effort and less expenses. It uses the functions of a standard mobile smart phone, owned by the visitor. Therefore no additional devices must be bought or maintained. The visitors can use their own smart phone and do not have to learn the function of a new device. No installations of activation sensors are required and the mobile smart phone in general works inside and outside of buildings. The mobile smart phone display gives the possibility to support the information representation with maps, pictures and animations. Therefore all technical criteria from the first section are fulfilled.

We include guidelines for human-computer-interaction and design different, multilingual and individual tours. The device handling is easy for the visitors, because they are familiar with their private mobile smart phones. Hence also all of the usage criteria from the first section are fulfilled.

Further, one major general novelty in our approach is the concept of using the visitor's personal mobile smart phone in combination with the manual activation by entering the appropriate object identification number and the digital watermarking based automatic activation. Our prototype application was rewarded with one of the Lowcost Multimodal Interface contest ("Loco Mummy Contest 2005") prizes, which was carried out by the EU Network of Excellence on Human-to-Computer Interaction, SIMILAR [17].

Future work will comprise the following aspects: Regarding the technical implementation, we have to develop a solution, how the audio comments can be played on additional types of mobile smart phones. This limitation we address in the moment, by implementing an adaptive audio decoder in the Audio Guide software.

The main focus of the future development is on the **human-computer-interaction**. Here we will concentrate more on the user adaptivity, for example with respect to users of different age groups. The goal is that each visitor, regardless how old, shall be able to use the Audio Guide application. Therefore the GUI (graphic user interface) must explain the usage of the Audio Guide software and should also be intuitive.

With current software implementation, we will perform an **evaluation** of the Audio Guide. For this reason we already recorded audio commentaries with content of the *"Meisterhäuser"* Dessau [1] to set up a pilot test of the Audio Guide for a limited duration at the *"Meisterhäuser"* realm. With the results of this evaluation we can then further develop and improve the Audio Guide software and the GUI.

Additionally, we will concentrate on evaluation of diverse **digital audio watermarking** parameters. Here we will focus mainly at robustness against D/A and A/D conversion and adding background noise, transparency, low complexity for during retrieval and the capacity for *OID*, *OID* with additional trigger flags and *OIW*.

Furthermore we plan to develop a **management and distribution platform** to support an easy update of content available to museums. Here on the one hand it will be possible to update the Audio Guide software. One the other hand the provider in the museum can enter new textual and audio information, define new exhibits for which the information

shall be available, delete exhibits or modify existing information, without knowing specific technical details of the application software.
Additionally, we will analyze possibilities to ensure Digital Rights Management (DRM) for the **DRM** of multimedia content in our application.
Finally, the **security aspect** of the Audio Guide shall be examined. Here we will identify which security problems can arise and evaluate solutions for the development of the Audio Guide software. The main aspects we already looked at are: Confidentiality, Authentication, Integrity, Availability and Non-Repudiation. For more details see the investigations in [18].

ACKNOWLEDGMENTS

This work has been partly supported by the city council of Dessau (Amt für Kultur, Tourismus und Sport), Germany and the EU Network of Excellence SIMILAR (Proposal Reference Number: FP6−507609). The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the European Union.

REFERENCES

1. *Meisterhäuser / Masters Houses in Dessau*, http://www.meisterhaeuser.de/, last requested 20.12.2005.
2. *Mediaexplorer... be closer*, http://www.mediaexplorer.ch/, last requested 07.09.2005.
3. *excerpts from EMP: The magical geek tour*, http://www.absoluterealtime.com/resume/meg.html, last requested 02.12.2004.
4. *Antenna Audio – the world is listening*, http://www.antennaaudio.com/xp.shtml, last requested 20.12.2005.
5. *eloqu – ebusiness.logistics*, eTour, http://www.eloqu.com/webpage_german/web_eloqu/mobil.htm, last requested 7.09.2005.
6. *Herzlich Willkommen auf den Internetseiten der cool IT GmbH*, http://www.coolit.ch, last requested 20.12.2005.
7. *BeyondGuide – Your personal audio guide on your mobile phone*, http://www.beyondguide.com, last requested 20.12.2005.
8. *Spatial Adventures – The foremost service provider of cell phone based audio tours and home of Mobile Touring Service*, www.spatialadventures.com, last requested 20.12.2005.
9. *User-Activated Audio Beacons (Ping!)*, http://www.touchgraphics.com/ping!.htm, last requested 21.01.2005.
10. *Rocky Mountain Audio Guide- We take you places*, http://www.rmaguides.com/, last requested 07.09.2005.
11. *Sennheiser, guidePORT – Exhibits come to life*, http://www.guideport.de//BRS_GuidePort_D_ 85059.pdf , last requested 08.09.2005.
12. *YOU HEAR WHAT YOU SEE - PICKUP AUDIO GUIDE*, http://www.dataton.com/pickup, last requested 20.12.2005
13. J. Dittmann, *Digitale Wasserzeichen: Grundlagen, Verfahren, Anwendungsgebiete*, Berlin 2000.
14. J. Dittmann, P. Wohlmacher, K. Nahrstedt, *Multimedia and Security – Using Cryptographic and Watermarking Algorithms*, IEEE MultiMedia, October-December 2001, Vol. 8, No. 4, pp. 54-65, ISSN 1070-986X, 2001.
15. T. Vogel, J. Dittmann, *Illustration Watermarking: An Object Based Approach for Digital Images*, In Proc. of Edward J. Delp III, Ping W. Wong (Eds.): Security, Steganography, and Watermarking of Multimedia Contents VII, SPIE Vol. 5681; SPIE and IS&T, pp. 578-589, Electronic Imaging Science and Technologie, 17-20 Jan. 2005, San Jose, California, USA,ISBN 0-8194-5654-3, 2005.
16. A. Lang, J. Dittmann, *Profiles for evaluation and their usage in audio wet*, in IS&T/SPIE's 18th Annual Symposium, Electronic Imaging 2006: Security and Watermarking of Multimedia Content VIII, Vol. 6072, P. W. Wong and E. J. Delp, eds., SPIE Proceedings, (San Jose, California USA), Jan. 2006.
17. *Loco Mummy Contest 2005*, http://www.locomummy.net/, last requested 20.12.2005.
18. S. Gebbensleben, J. Dittmann, C. Vielhauer, *Sicherheitsaspekte eines Mobilen Multimedia User Guides*, accepted paper on conference D-A-CH Security 2006, 28-29 March 2006, Düsseldorf, Germany