ABSTRACT

The design of immersive soundscape experiences, both for artistic and informative purposes, is an established field in Auditory Display. This paper describes the process of designing historically informed soundscapes to be incorporated in modern travel-guide applications. The work stems from the research project TRACCE (TRavologue with Augmented Cultural & Contemporary Experiences), which focuses on the design and development of a platform for augmented cultural routes. Using this platform, hikers can follow the journey of 18th and 19th century travelers, having access to the original travelogues, tracked routes, and a wide variety of modern information. User experience is achieved by means of visual and auditory reconstructions of the original surrounding environments in several identified points of interest in each path. Apart from the creative process and technical details, the paper discusses the design challenges, which mainly stem from a) the limited data available which would allow an accurate and convincing reconstruction of the acoustic environments, b) the need for diverse auditory displays which would grasp the users’ attention, and c) the difficulty in designing soundscapes which would be interesting, appropriate, and informative for a wide audience of various age groups, educational backgrounds, and sensory abilities.

1. INTRODUCTION

Since its original introduction in 1969 by Murray Schafer [1], the term Soundscape has grown to encompass various scientific and creative aspects. It relates to the early studies on Acoustic Ecology, which sprang from the realization that people had limited awareness of their surrounding acoustic environment and aimed at increasing public awareness of the human impact on it. In a recent definition, the term soundscape is used to describe the entire acoustic energy present in a landscape (all foreground and background sounds), relative to the orientation and position of the listener in that space. It is further argued that foreground sounds, which are either caused by or require the immediate attention of the listener, are the least associated with the landscape, in contrast to background sounds, which are low-level events from multiple sound sources, blended into an indistinguishable body, that are highly associated with the surrounding landscape [2].

This definition acknowledges the presence of various levels of auditory events in soundscapes, fact which has long been implemented in the work of various scientists and artists. Modern soundscape compositions are associated with various auditory phenomena, ranging from sounds of the nature to artistic practice [3]. Their goal is to invoke memories and associations related to a familiar physical space while stimulating audience imagination.

Audio-walks, guided-tours, and virtual-audio tours are example applications under the soundscape design and composition domain. Audio-walks and guided tours are outdoor or indoor explorations, which are augmented by pre-recorded information, such as sound-field recordings, narration, music, or any combination of the above, usually played back via headphones for each user. They are becoming a ubiquitous way to experience history and heritage of any form, and are often associated with visitor experience in museums, archaeological sites, galleries, etc. [4]. They hold an advantage over other sources of information, such as written signs, as they overlay an auditory memory of a given place to its present sound-field, creating a powerful linkage [5, 6]. Their aim is to enhance visitor experience by focusing on every aspect of the soundscape thus drawing people out of the visual overload and into the sonification of their walk through the area of interest [4].

Virtual-audio tours offer users the possibility to virtually travel to any place in the world and experience the unique acoustic characteristics of different soundscapes in an immersive manner from the convenience of their homes [8, 9, 10]. They heavily rely on field-recordings, which capture the core characteristics of soundscapes at specific points in time. These auditory “documentations” of different locations are then used to stimulate listener-interest, creating a form of virtual sound tourism.

This paper presents the process of designing historically informed soundscapes, to be incorporated in modern travel-guide applications. The creative process and technical details are being described, followed by a discussion of the design challenges. These challenges result from the lack of data that is descriptive of the sound qualities of soundscapes which people experienced 2 centuries ago, the need for diverse auditory designs captivating the users’ attention, and the difficulties in designing appropriate and informative content for a wide audience of various age groups, educational backgrounds, and sensory abilities.
2. TRACCE

TRACCE (TRavelogue with Augmented Cultural & Contemporary Experience) is a multi-partner nationally-funded research project, which aims to design and develop an innovative travel platform for mobile devices (iOS and Android) featuring cultural routes. The services address mainly to tourism professionals and cultural institutions, in order to design specialized tours (routes, guided tours, exhibition guides, interpreting programs, etc.) that will enhance visitor experience. Within the TRACCE project the functionality of the platform is indicated through the creation of several routes of increased cultural interest, based on historical and ethnographic content provided by the Laskaridis Foundation in Greece.

Using the proposed platform, users will be guided through the historical trails of travelers, who visited certain parts of Greece in the 18\textsuperscript{th} and 19\textsuperscript{th} century, having access to the narrative pattern of their journeys (travelogue literature), while maintaining access to a variety of modern information about each route. In this way, a user’s experience of the physical visit to an area of interest is uniquely enriched by the inevitable comparison between historically documented and contemporary touring data logs, combined at will and in a seamless manner. The narrative output of the system is updated in real-time to adapt to the route choices of the user during the physical tours. Its contents can be enriched with virtual reconstructions of the natural landscape using the camera and screen of the mobile device for viewing digital material as well as with the audition of historically informed soundscapes, which, when presented in conjunction with the physical environment, are expected to augment user experience.

Part of the content created within the TRACCE project will be made available free of charge, aiming at increasing the touristic visibility of different parts in Greece, consolidating cultural routes as a brand that enhances visitor attraction, and introducing new business opportunities for activities in these regions. The final platform will feature three paths: Crete, Northern Greece, Ionian Islands and the soundscapes content will be available to users in two ways: off-site, via pre-downloading, or on-site, via audio streaming, using a cellular or WIFI connection.

3. HISTORICALLY INFORMED SOUNDSCAPE DESIGNS

Within the scope of the TRACCE project, the Laboratory of Music Acoustics and Technology (LabMAT), of the National and Kapodistrian University of Athens, is concerned with research on the design of historically informed soundscapes and their evaluation. The main sources of information for these designs are travelogue texts from the 18\textsuperscript{th} and 19\textsuperscript{th} centuries (Figure 1) and relevant iconographic material from the same era (Figure 2).

The cultural routes follow a path of pre-defined points of interest, the majority of which were selected based on the travelogues. Some additional points were added to include modern areas of interest along these paths. Several route options were drawn from each travelogue, in an attempt to reach the interest of a wide range of audience. Some additional points were added to include modern areas of interest along these paths. Several route options were drawn from each travelogue, in an attempt to reach the interest of a wide range of audience. For all identified points the following metadata has been catalogued by our collaborating partners to assist navigation of the content: reason for travel (e.g. politics, religion, archaeology etc.) date of travel, means of transport, location name (past and modern), location co-ordinates, accommodation, time and duration of transport, relevant pages on the travelogue, comments regarding the people the writer met.

Extensive research is necessary to design historically informed soundscapes. That is soundscapes designed to match as closely as possible the soundfields experienced by the travelers in those specific physical spaces during a specific time in the past. The main sources of information for the acoustic ecology of these places al-

https://tracceproject.eu/en
most 200 years ago, are the descriptions of the scenery in the travelogues themselves. The relevant pages to each point of interest were carefully read in order to detect any direct mentions of nature or human driven sounds. More often than not, auditory information was indirect and potential sound sources were inferred from the writer’s general description of their surroundings.

The “auditory puzzles” were complemented with information extracted from the drawings and sketches found in the texts and the general iconographic repository of the project, as well as modern cartography. We have relied heavily on the available visual information of each location to draw conclusions on the morphology of the landscape in order to place sound-sources on meaningful and accurate positions in space, and to simulate the acoustic characteristics of each place as closely as possible to reality. When necessary, further information was drawn from additional resources, either contemporary to the travelogues or newer material referring to the same era.

The designed soundscapes are fixed media compositions comprising of newly conducted audio recordings and audio samples from public libraries, combined in such a way to describe selected excerpts from the travelogues. The content is augmented with narrations of short excerpts from the original texts, intended to help listeners understand the story that unfolds around each point of interest, and abstract musical elements designed to enrich user-
experience.

As it will be detailed in the next Section, all sound elements were mixed and grouped in four stems. Each of these elements and stems were processed creatively and, when necessary, 3D spatialized to create a more immersive experience for the user. At the current stage of the project, no attempt was made to either personalize the utilized Head-Related Transfer Functions (HRTFs) nor account for the audition conditions (headphone based, speakerphone, real-world sound interference etc.). The potential impact of these choices on user immersion is discussed in Section 5.

Within this framework, the aim is to achieve a structural and sonic balance between the various components of each soundscape. This leads to sound designs that are “loyal” to the space acoustics, audio spatialization, and sound content, and can captivate the listeners. The resulting soundscapes are never longer than 90 seconds in keeping up with the rule of thumb that the maximum duration humans can remain concentrated on a listening activity is less than 3 minutes [5].

4. SOUNDSCAPE ARCHITECTURE

Similar to museums, science/cultural centers, zoos, etc., which concern the education and entertainment of a very diverse visitor population, the platform fostering the guided augmented cultural routes of the TRACCE project falls within the scope of Informal Learning Environments (ILEs) [12]. As such, the soundscape content, created to augment this experience, must be interesting, appropriate, and informative for a wide audience of various age groups, educational backgrounds, and sensory abilities.

In an attempt to address this issue, the content of the designed soundscapes was divided into 4 layers of different audio event-types: Background, Storytelling, Narration, and Music. Each of these layers can exist independently but, at the same time, they can all be played back in any combination. Users have complete freedom to control the content of the soundscape they hear, by enabling and changing the relative level of the corresponding layers(s) in any combination they desire. All soundscapes compositions follow a first-person perspective, that of the travelogue writer. The Narration and Background layer audio tracks are monaural and stereophonic, respectively, while those of Storytelling and Music are binaural, enriching the immersive experience of the users.

This approach of multi-layered soundscape designs builds upon previous research. There have been documentations of various soundscape designs combining pre-recorded text, music, and ambient environment sounds with the real-time original ambient sound of the surrounding space of audition [13][14]. Various terminologies have been proposed to categorize the auditory information present in soundscape designs with similar taxonomies to the one discussed here, such as: biophony (sounds of animals, birds etc.), anthrophony (sounds produced by humans), and geophony (sounds of the atmosphere) [15], or keynote sounds (ambient environment sounds), sound signals (foreground/focus sounds), and soundmarks (sounds unique to a soundscape) [16]. Regardless of the terminology used, the need for such taxonomies implies that soundscapes are complex auditory entities whose components should be handled differently according to the information load they introduce. As explained in detail below, each layer in the soundscapes designed for the TRACCE project contains different types of auditory content and carries different levels of information. Hence, with the selection and level adjustment options offered, users have the ability to create personalized versions of the soundscapes that best fit their needs and preference.

The advantages of auditory display designs combining narration, non-speech audio, and music have also been debated in the literature. The use of narratives, that are brief enough so as not to restrict but rather guide the auditory experience, have become increasingly popular in audiowalks. The ability to listen to another person’s story, experience, or memory referring to the exact same location the user is standing can cause moving emotions, and lead users experience their surroundings differently [5]. In the context of educational experiences, it has also been argued that narratives provide a rich, entertaining, and fully functional means for the communication of information [17].

The use of music or musical excerpts, either pre-composed or sonified based on real-time data, in guided walks has also been encountered in recent studies and applications. In this context, music adapts dynamically to the users’ location and movement using GPS data from the users’ smartphones. Such musical layers in complex soundscape designs have been reported to shape user experience, and enhance their engagement to the composite auditory content [13].

The following paragraphs discuss an example soundscape scenario and use its realization to describe the technical details of its 4 layers.

4.1. Example soundscape scenario

The selected example originates from the first chapter of Volume I of the “Travels in Crete” travelogue [18] by Robert Pashley (1805-1859), a British economist who travelled around Crete in the early 19th century, back when the island was still part of the Ottoman empire. It was the 8th of February, 1834, when Pashley arrived at the gulf of Kania (sic). The travelogue starts with a mention of the view of the White Mountains, which seem to have snow on them, and the minarets of the mosques that emerge above the rest of the buildings in the city. Upon arrival by a small boat into the harbor, Parsley and his companions are met by a Turkish and a British Consul, with the latter insisting on having him as his guest. The text fast forwards to the sunset, when a salute is fired from the guns of the fortress.

4.2. Background

The Background layer comprises sounds that attempt to reconstruct the auditory environment of the landscape, such as sounds of nature (wind, air) and content that is considered background noise to the travelogue writer. The audio track is stereophonic, with left / right panned stationary and moving sources. This layer makes extensive use of field recordings which serve as the ambient base for the historical reconstruction of each soundscape. These recordings were originally planned to take place at the physical locations described in the travelogues. Yet, because of the travel restrictions posed due to the COVID-19 global pandemic, they were substituted by recordings in rural locations of equivalent acoustic characteristics around Athens. The potential impact of this choice is discussed in Section 5.

In the soundscape corresponding to the aforementioned scenario, the Background layer reconstructs the audio environment of the arrival at the port of Chania as well as the disembarkation from the ship. Taking into account the descriptions in the text and the seasonal characteristics, one can hear the sound of the air and the splash of waves on the moving ship along with the sea sound on the
boat in which the writer disembarks at the port. The sound of the sea and air change from the open sea to the port complete the audio scenery. Upon disembarkation, the Background layer changes to reflect the crowded port surroundings.

4.3. Storytelling

The Storytelling layer consists of a binaural audio file, which includes sounds that are caused by the writer's interaction with the surrounding area (e.g. his own footsteps as he navigates the space while the story is unfolding), as well as sounds caused by other factors which have a direct relation to the action of the story as it appears on the text (e.g. shotguns). In this way the user of the application revives in a first person view, the writer's auditory experience in an immersive manner. Sound spatialization is implemented using the Anaglyph binaural engine [18].

In the Chania soundscape example, one can hear sounds from the ship funnel and anchor, the lowering of the small boat, which transfers the passengers from the ship to the shore, into the sea, and its oars. During disembarkation, the layer contains the steps of the passengers leaving the boat and walking through the port, sounds from a minaret afar, and the voice of its Imam calling the locals to prayer. The clip concludes with human-generated sounds from other people around the port and steps as the traveler leaves the port premises.

4.4. Narration

The Narration layer contains a monaural audio track consistently panned to the center, which includes selected narrated excerpts from the original text that are directly related to a specific point of interest, and are difficult or impossible to render clearly through the auditory reconstruction of the soundscape. Fragmented narration is considered more effective than a whole story, allowing users to better connect with the location while structuring a cinematic approach of a given scene thus creating a more vivid impression of the story [5].

In the discussed example, the text includes the following excerpts from Pasley's travelogue: “On entering the gulf of Khania I was struck with the grandeur and beauty of the White Mountains. [...] As the boats of the Hind pulled into the harbor, to land me with my companions, we were asked, in a language the sounds of which I had not heard for several months, whether we had come from a Turkish port. [...] At sunset a salute was fired from the guns of the fortress, [...] During my stay at Khania I became acquainted with its inhabitants” [11].

4.5. Music

The Music layer contains a composition of electroacoustic music realized with the stochastic synthesizer CosmosF, which is inspired by Xenakis formalized music [19] [20]. CosmosF takes as input audio files which can be manipulated in real-time by controlling macro, meso, and micro time-scale parameters. A similar parameterization and stochastic approach has been employed at the Tettix project which deals with the analysis and modeling of cicadas’ singing, using as data soundfield recordings of cicadas in archaeological sites around Greece [21]. The source material (input audio files) used in the TRACCE project are the Background, Storytelling, and Narration layers. Stated differently, the source material for the music compositions are stochastic interpretations of the remaining 3 layers of the soundscapes. In this way music blends smoothly with the remaining content while maintaining its unique characteristics. Similar to the Storytelling layer, Music is also enhanced by the use of 3D audio spatialization. Spatialization is once again introduced through the Anaglyph binaural engine [18].

5. DISCUSSION

While the term soundscape was originally strictly associated with ecology, it has rapidly evolved to encompass various interdisciplinary fields, from purely scientific to artistic. Especially for applications with a didactic focus, the use of soundscape compositions, augmented with narration and music, have been found to be particularly effective [17][22].

Nevertheless, a potential weakness in the practice of on-site field-recording audits has been raised in the literature, concerning the fact that the immersive illusion these audits create is fragile and prone to sudden disruptions, when the pre-recorded content contradicts the live visual and auditory input [22][24]. Yet, should the TRACCE application users choose to listen to the audio content on site, it is precisely this momentary destabilizing effect created by the direct comparison of the contemporary to the historical soundscape that is intended. Users are expected to be confronted with variations and similarities between the reconstructed soundscapes and the surrounding sound-fields, as a means of experiencing the evolution of the landscape, as a result of human action. After all, according to Hogg [24], a soundscape consists not only of its contemporary sonic state but also of its past content, which includes human and nature induced sounds. Seeing it from this perspective, these "condensed" (90 second) auditory field-recording auditions has been raised in the literature, concern...
form of virtual sound-tourism.

Our original intention was to embed in the designed soundscapes, field-recordings collected on-site at the points of audition. These recordings were supposed to form the basis for the Background layer, in order to be in accordance with the landscape. Nevertheless, travel restrictions posed due to the global COVID-19 pandemic prevented this intention from being realized. As an alternative, field recordings were conducted in rural locations of similar landscape characteristics around Athens. Upon collection these recordings were carefully auditioned and stripped of any occasional sound events which could not possibly belong to the era of the travelogues (e.g. sounds of motor vehicles). Similarly any audio excerpts which included speaking voices were also removed. This action was necessary as spoken language has obviously evolved over the centuries, and the inclusion of “modern-Greek” audio excerpts in the soundscapes would render their historical accuracy invalid. The impact of the choice to use audio recordings from different locations on the coherence between the auditory and the visual content at the actual points of audition, remains to be evaluated through on-site subjective studies, once the conditions permit it. Nevertheless, since the designed soundscapes intended to approximate the soundscape conditions of 1 to 2 centuries ago, a certain level of disassociation between the landscape and the soundscape is expected and to a certain degree intended. Hence, it is hypothesized that the selected alternative will not have a significant impact on user experience and immersion.

Moreover, in the current iteration of the project the Storytelling and Music layers are spatialized using a generic HRTF set. An investigation of the impact of HRTF personalization and selection on content appreciation and immersion, in the context of such an application, is in the immediate plans of the authors. The necessary level of immersion for this type of applications, one which would allow users to experience an auditory “time-travel”, while maintaining awareness of their surrounding soundscape, remains to be defined. Such an investigation should not ignore the uncontrolled audition conditions (headphones, ear-buds, noise-canceling headphones, speaker-phone). Similar concerns have been raised in other soundscape composition projects, where subjective evaluations indicated that users did find the level of realism, provided by non-personalized HRTFs, to be acceptable [26]. In the context of the TRACCE project, the Anaglyph binaural engine was chosen because it offers several options for HRTF selection / personalization and could easily foster such an investigation.

Another point that needs to be addressed concerns the potential impact of audio compression on spatialization. Since the soundscape content is intended to be streamed to the users’ mobile devices for on-site audition, some form of compression may have to be applied on the audio content to facilitate this process. Its potential impact on user immersion remains to be evaluated through ecologically valid user testing. As discussed above the goal of these studies will be to identify the level of immersion that is realistic and necessary in such unstable / unpredictable audition conditions so as to adjust all audio-related parameters accordingly.

The effectiveness of the designed historically informed soundscapes in augmenting user experience within the context of cultural routes are in the immediate plans of the authors and will be realized both on-site and off-site once the conditions permit it, using the designed travelogue app. It is also in the authors immediate plans to investigate the level of necessary spatialization accuracy for user immersion, through a series of ecologically valid subjective studies, taking into consideration the audition conditions and the type of content spatialized.

6. ACKNOWLEDGMENT

This research has been cofinanced by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH CREATE INNOVATE (project code:T1EDK-02146).

7. REFERENCES


