



**MÄLARDALENS HÖGSKOLA
ESKILSTUNA VÄSTERÅS**

Thesis Advanced Level, Computer Science CDT504 (30.0 hp)

Master's in software engineering | 120.0 hp

Music in Motion - Smart Soundscapes

Usman Alam

uam10001@student.mdh.se

Supervisor

Jennie Andersson Schaeffer

Co-supervisor

Afshin Ameri

Examiner

Baran Çürüklü

MÄLARDALEN UNIVERSITY

School of Innovation, Design and Engineering

June, 2018

Table of Contents

<i>List of Figures</i>	4
<i>List of Tables</i>	5
<i>Abbreviations</i>	6
<i>Acknowledgements</i>	7
<i>Sammanfattning</i>	9
<i>Abstract</i>	11
<i>Chapter 1: Introduction</i>	13
1.1 Background	13
1.2 Problem Statement	14
1.3 Research Objectives	14
1.4 Motivation.....	15
1.5 Research Outline	15
<i>Chapter 2: Literature Review</i>	16
2.1 Introduction.....	16
2.2 Related Work	16
2.3 Theoretical Framework.....	18
2.4 Multimodality	19
2.5 Communication in Multimodal Theory	20
2.5.1 Discourse.....	20
2.5.2 Design	21
2.5.3 Production.....	21
2.6 Research Gap	22
<i>Chapter 3: Research Methodology</i>	23
3.1 Introduction.....	23
3.2 Iterative Design Process.....	23
3.3 Data Collection (Primary and Secondary Data).....	24
3.4 Sampling Approach	25
3.5 Data Classification and Analysis	27
3.6 Design and Development	31
3.7 Author’s Role and Contribution.....	32
3.8 Research Ethics.....	33

Chapter 4: Results and Discussion	34
4.1 Introduction.....	34
4.2 System Overview	34
4.3 Translating Multimodal Design Brief (MDB)	35
4.3.1 Translating Scenario “Provoke by Standing Visitor”	35
4.3.2 Translating Scenario “Change in Interaction”	37
4.3.3 Translating Scenario “More than Six Visitors”	40
4.3.4 Translating Scenario “Based on Distance”	42
4.4 Functional Requirements Implementation and Graphical User Interface	42
4.5 Discussion	46
Chapter 5: Conclusion	49
5.1 Future Work and Suggestions	50
References	52

List of Figures

<i>Figure 1: Research process Bill Wolfson, 2013</i>	<i>24</i>
<i>Figure 2: Stockholm history museum artifact chosen for project demonstration.....</i>	<i>25</i>
<i>Figure 3: Gesture drawing to express different meanings of body movements</i>	<i>28</i>
<i>Figure 4: A segmentation of a sample interaction.....</i>	<i>29</i>
<i>Figure 5: A segmentation of Albatross and Duck dance gesture.....</i>	<i>29</i>
<i>Figure 6: An abstract sketch of interactive environment.....</i>	<i>34</i>
<i>Figure 7: Relations between sounds provoked by standing visitor.....</i>	<i>37</i>
<i>Figure 8: The change in sound by changing in body position.....</i>	<i>40</i>
<i>Figure 9: Play sounds for more than 6 people in soundscape.....</i>	<i>41</i>
<i>Figure 10: System's main menu screen</i>	<i>43</i>
<i>Figure 11: Screen which facilitates to link sounds</i>	<i>44</i>
<i>Figure 12: For adding/setting up relations</i>	<i>44</i>
<i>Figure 13: Simulator for testing generated soundscape.....</i>	<i>45</i>
<i>Figure 14: View sounds relationships</i>	<i>46</i>

List of Tables

<i>Table 1: Abbreviations</i>	6
<i>Table 2: Observations collected from museum</i>	26
<i>Table 3: A list of interactions, selected by us from the collected data and suggested by students.</i>	30
<i>Table 4: Sounds provoked by standing visitors</i>	35
<i>Table 5: Change in movements and change in sounds</i>	38
<i>Table 6: More than six-person scenario</i>	40
<i>Table 7: Other sounds</i>	42

Abbreviations

Table 1 describes abbreviations used in this thesis work.

Table 1: Abbreviations

Abbreviation	Meaning
ACM	Association for Computing Machinery
API	Application Programming Interface
BLL	Business Logic Layer
DAL	Database Access Layer
FCL	Framework Class Library
GUI	Graphical User Interface
HCI	Human-Computer Interaction
IEEE	Institute of Electrical and Electronics Engineers
MDB	Multimodal Design Brief
MDH	Mälardalen University
MiM	Music in Motion
MS	Microsoft
ROS	Responsive Open Space
RQ	Research Question
SQL	Structured Query Language
SRS	Software Requirement Specifications

Acknowledgements

My particular thanks to the thesis supervisor Jennie Andersson Schaeffer and co-supervisor Afshin Ameri. I am grateful for their invaluable constructive criticism, aspiring guidance and many insightful conversations during this work.

I also express my warm thanks to my thesis examiner Baran Çürüklü, senior lecturer, Mälardalen University, for sharing his honest views and suggestions on many issues related to this work.

I would like to extend gratitude to Ms. Hanna Petersen and Ms. Klara Nordquist, students, Mälardalen University. The multimodal design brief was their idea which helped to implement the smart soundscape system.

Most importantly I would like to thank Mälardalen University who supported the project “Music in Motion - Smart Soundscapes” and provided me an opportunity to contribute in the research area of multimodal interaction design.

Lastly, I would like to thank my family, friends and my colleagues at Awave AB for their love and continuous support.

Sammanfattning

Denna studie är en del projektet Music in Motion (MiM). Projektet genomfördes i samarbete mellan Mälardalens högskola och Svenska Historiska Museet i Stockholm. Det syftar till att utveckla ett system som kan generera ljud baserat på människors interaktion med systemet eller miljön. Det ska kunna installeras olika artefakter val.

Avhandlingen fokuserar på att designa och utveckla ett system som heter Smart Ljud System. Kärnan i detta system är baserat på utforma en algoritm som kan tolka besökare beteenden och deras sätt att interaktion med artefakten och; tilldela ett unikt ljud för varje interaktion. För att utforma ett sådant system är det viktigt att undersöka de parametrar som helst kan representera människors interaktioner och hur de kan påverka de ljud som genereras. I denna uppsats tar interaktioner plats i en museimiljö. De primära data samlades in genom att observera samspelet mellan besökarna och en artefakt i museet.

I denna studie beskrivs också hur man kan bryta ner en idé som heter Multimodal Design Brief för utveckling av ett smart ljudsystem. Det första steget var att utforma de funktionella kraven genom förståelse av den multimodala design briefen och efteråt anpassa systemet för dem. Denna avhandling följde en iterativ designprocess metodik. Genom att använda denna metod har målen för denna studie uppnåtts och en fungerande version av systemet utformas och genomförs.

Nyckelord: *Multimodalitet, HCI, Ljudlandskap, Digitala museer, ljudprogram, Interaktion och Interaktionsdesign*

Abstract

This thesis aims to explore the possibility of developing a system that can interact with people near it, through composing and playing unique tailored sounds for them. This research is a part of the Music in Motion (MiM) project, which was carried out in collaboration between Mälardalen University and the Swedish History Museum in Stockholm. It aims to develop a system which can interact with people and generate soundscape based on the people's interaction with the system or environment. It can be installed on any artifact of choice.

The thesis focuses on designing and developing a system named Smart Soundscape System. The system is based on designing an algorithm which can interpret visitor's behaviors and their way of interaction with a museum artifact; therefore, allocate a unique sound to each interaction. To design such systems, it is important to investigate the parameters which can ideally represent people's interactions and how they can affect the generated soundscape. In this thesis, the interactions are considered in the context of the museum. The primary data was collected by visiting and observing the interaction of visitors around an artifact in the museum.

In this study we give an interactive environment scheme that helps to explore the possibilities for generating interactive space, in which visitors and a museum artifact become involved in collaboration. Moreover, the author shows how this scheme can be defined to extract the functional requirements and implement them for the smart soundscape system. This thesis followed an iterative design methodology. By using this method, a functional version of the system is designed and implemented.

Keywords: Embedded systems, Multimodal, Gesture technologies, Human-computer interaction, Soundscape, Digital museum, Audio applications, Interaction, Interaction design

Chapter 1: Introduction

1.1 Background

Museums are vital source of spreading culture and history, as they exhibit different objects which are essential to any culture. As we are living in the digital age, where mobile and computer devices are becoming increasingly dominant part of our daily lives. Similarly, virtual lives are becoming more common with personalized user experiences. In this context, museums have to transform according to the change in digitalization and discover new approaches to tell stories about exhibits. This will certainly help to the museums to attract more audiences and increase their interests [3].

During past two decades, the shifts in digital technology and society already gave a new orientation to the museums. In other words, to provide immersive and collaborative experiences for audiences many museums are using technologies such as digital services (includes social media etc.), tablets, and sensors for smart environments. In future, more and more changes in technologies will put pressure on museums to innovate contemporary trends to meet people expectations. They will continue to be reshaped by numerous trends and technologies to attract broader audiences [3].

Museums are becoming more collaborative; they can utilize data generated by visitors through a collaboration of mobile devices and sensors. Data collection and analytics combined with automation will make a feasible way for them to deliver fascinating individual experiences where visitors, museum artifacts, sound, light, and space interact to generate a smart environment. This strategy enables the museums to constantly develop and increase their exhibitions, and immersive experiences [3].

Museums are designing interactive exhibits and making use of physical interactive installations to increase visitors' engagement. Contemporary museums are compliant to technologies which can bring transformation in the learning process. By using such technologies efficiently, we can provide a refreshing and exciting experience of interactions among visitors and museums objects [27]. The term interaction is used to deliver an abstract idea which can be widely used in different fields. In general, the concept of interaction is an exchange of information between two or more entities [13].

Information exchange can happen through various channels or modalities. Multimodal interaction provides a platform for users to interact with the system through multiple modes such as speech, touch, gestures and body movements. Since such interactions are natural to humans, multimodal interaction is becoming popular for culture-related public spaces such as experience centers, festivals, and museums [20]. A museum is a good place for exploring the possibilities of using digital domain to attract the audience; so, designing interactive installations for museums have become quite common [13].

The aim of this thesis is to design and develop a smart soundscape system that generates soundscape based on interactions between visitors and a museum artifact. The smart soundscape system is part of the project Music in Motion (MiM) at Mälardalen University, Sweden. MiM was carried out in collaboration with the Swedish History Museum, Stockholm. The primary goal of the MiM project is to develop a system that can interact with people through playing unique tailored sounds. The system will also have the capability to recognize interactions with several people and assign different sounds to each of individual. Here, people will be able to generate a soundscape in collaboration. The system can be installed at various places such as museums, train or bus stations, public parks, airports, etc. To achieve the research aim the first step is to formulate the research question to achieve the research objectives, accordingly.

1.2 Problem Statement

The purpose of the thesis is to investigate and develop a system that can map the visitor's behavior and interactions. And generate different sounds with the change in environment. A general question comes into the mind.

“How can natural movements of museum visitors be used in the generation of soundscape?”

Sensors and computer devices are becoming more and more consolidated in our everyday lives. We can create a useful smart environment in conjunction with sensors around museum's artifacts. This phenomenon can create a seamless environment for visitors where they can interact with exhibit and system for the generation of soundscape. Besides, it can also enable the visitors to act upon sounds in the interactive space.

One of the challenges is to design such interactive environment which enables the visitors to interact and engage with the museum artifact, using gestures and natural movements. It will not only address the needs of smart soundscape system but also help to collect the requirements to implement the system. As the artistic work is an incremental process and the artist might need to revise the design several times. Therefore, the author focuses on a rapid way to translate those designs into technical requirements. Based on this rationale, the research question that arises for the thesis is:

Research Question (RQ): *How to break down an interactive environment scheme into functional requirements, so that they can be implemented for smart soundscape system?*

1.3 Research Objectives

The objectives of the research are as follows:

- Investigate what parameters can best represent the visitor's interaction and how these can affect the generated soundscape.
- Research to determine how an interactive environment scheme can be designed to better recognize interaction and assist the system.
- Investigate the methods that can be useful for translating an interactive environment scheme to a computer program.

- Seek the possibility of developing a system that can interact with people near it, through composing and playing unique soundscapes for them.
- Developing a system with a capability to recognize interactions with several people and assign different sounds to everyone.
- Making the system convenient enough to be installed at various places such as museums, train or bus stations, public parks, airports, etc.

1.4 Motivation

This section discusses aspects that persuade the research undertaken in this thesis. The motivation behind this research is to combine the arts with technologies. To the best of our knowledge, there is very little work done in this context. The reason might be that there is negligence about arts and technology. This lead us to an investigation of combining the arts with science. This investigation may provide some design ideas which can influence the user experience and design of future public spaces, such as museums, parks, airports etc. We are interested in designing a smart environment where exhibits, visitors, space and sound collaborate with each other to create a soundscape together. To achieve this, we decided to explore the possibility of developing a smart soundscape system. It is argued that public places are already practicing by installing such systems. However, this researcher is particularly about mapping the visitor's behavior and interaction and generate a tailored sound for the visitors. Therefore, we choose this topic for research purpose and check its feasibility while developing the system.

1.5 Research Outline

The rest of the thesis is structured as follows.

- The section 1.1 of Chapter 1 provides an overview of the research topic. Section 1.2 describes the problem statement for the research and section 1.3 describes the research aims and objectives.
- Chapter 2 summarizes a multimodal design used for interactive environment and theoretical framework related to multimodality (Section 2.3). Further on, section 2.4 describes the theoretical analysis of communication in multimodal theory discourse, design, and produce relevant to this research.
- Chapter 3 focuses on requirements gathering, research methodology, design research, data collection and analysis, and development methodology of the thesis.
- Chapter 4 provides the results of this thesis (Section 4.2) and implementation of smart soundscape system (Section 4.3). It includes a detailed explanation of critical discussion about this thesis.
- Further on, the last Chapter 5 contains the conclusion (Section 5.1), contribution to the literature in Section 5.2. This chapter presents the possible suggestions for future improvements in the system (Section 5.3).

Chapter 2: Literature Review

2.1 Introduction

This chapter describes the theoretical framework for the research topic. This work focuses on the requirements of developing a system which can generate a soundscape based on the interactions between visitors and a museum artifact. In this context, the input to the system depends on gestures; and the output is a tailored sound. The system comprises of two sides: the human and the system. The human side is defined by the body interaction of the visitors who are present in the soundscape. The body interactions are in the form of gestures that provide sensory input to the system. The system side is focused on an algorithm which can be implemented to play tailored sounds by the input. In other words, the system should assign different sounds to each visitor when there are several visitors interacting with it. In this manner, visitors can team up with each other to create a soundscape together.

The next section discusses the related work to this thesis and explains their potential similarities and differences.

2.2 Related Work

Gau, et al., 2012 in their work [11] the technical part of their system includes four components. The author will discuss two components (i.e. *tracking system and generative system*), as they are related to this work. The goal of the *generative system* to generate soundscape and provides relative feedback corresponding to the contributor's activities and body movement in the space. One of the aims of their system is, to detect how many participants are present, measure their distance to each other, their position and movement orders. Their system then transposes the measured data in real time into several musical parameters such as sound pitch, period, rhythm, harmonics etc. These parameters could affect soundscape density. Whereas, our system uses different parameters such as the number of visitors, their walking speed or interactions with the system to create a tailored melody for the environment. However, the smart soundscape system will take these parameters as inputs from an integrated tracking system named as Kinect system. So, the medium for receiving information is not the same as in our work. In addition, the Kinect system provides 360 View. 360 View allows visitors to send combination of body movements and interact with the system without holding a control unit in their hands. In contrast, their [11] system does not offer such view.

Furthermore, in Gau, et al work [11] each participant's interaction is detected by tracking component of the system and predefined sounds are assigned to each participant. Likewise, in our approach the system will recognize when there are several people interacting with the system and assign different sounds to each visitor, this way visitors can collaborate with each other to create a soundscape together.

Gau, et al., 2012 work [11] suggested a *performative* dimensional environment named as *Responsive Open Space* (ROS). The ROS facilitates collaboration between participants by combining audio-visual compositions. The technical system specification includes “*multi-user tracking system, architectural system of surfaces, space responsive sound system and the real-time visualization*” [11]. Further on, the ROS targets the human to the human relationships and allows the contributors to generate their individual “*performative space*”. In this context, the ROS does not investigate the relation between human and the machine. Therefore, in their work [11] for the interconnectivity of participants and interaction among them, at least two or more people are required as they did not concentrate on human to machine interaction. On the other side, this thesis focuses on human to machine interaction as well. Hence, at least one human is required to activate the system interaction.

For communication of art in the physical museum space, Karen and Kaj [19] provided insights into the innovative approaches to interaction design. In their work, they discuss four approaches and how they implemented installations for museum space based on these methods. One of them (i.e., the sound of art installations) is relevant to this thesis. In their work “The Sound of Art Installations,” they defined 25 different spots in a museum. Each spot is drawn visually on the floor with silver color and a passive infrared sensor which is directed towards a speaker that is present above the circle. When a visitor enters the loop, the sensor registers visitor’s movements inside that spot and triggers to play an audio clip. The visitor can only hear the sound independently while standing inside a circle. The audio sounds are randomly selected and if the visitor makes a tiny movement, a new sound is played until all the sounds have been played an equal amount of times [19].

Similarities with this work can be identified in “the sound of art installations” approach, as Karen and Kaj [19] are considering visitor’s movements in defined spots through sensors and playing audio sounds. However, in their approach, they focused on independent user experience for hearing the sound; while, we will provide a shared experience for the visitors as well. One more difference is that they are selecting sounds randomly to play, whereas, this project would focus on playing sounds according to specific interactions as defined in the system. Lucia and Andreas (2004) presented the key concept that for a broad user experience, the individual perception of visual, auditory and imaginary space and their relationships should be placed at the center of the interface design. They focused on the issue in connection with multimodal interaction, by taking the sense of hearing principles [22].

The sounds can be experienced by visitors in open space loudspeaker. In their project LISTEN, they have developed a tailored immersive audio argument surrounding for the visitors of art exhibitions. The museum’s visitors can experience the customized sounds about exhibits by wearing wireless headphones [22]. On the other side, our proposed system focuses on multimodal interaction by taking the body movements of visitors, i.e. gesture, touch, and voice. Thus, the input to smart soundscape system is based on gesture and output is tailored sound. Besides, another

property in which their work differs from our is that LISTEN provides a personalized immersive augmented environment.

Astrid, et al., 2012 [4] in their work “*KlangReise*” tried to make visitors aware of naturally occurring soundscapes by installing the system in some museums. They observed that the visitors’ responses for further investigation, experimentation, and the development for soundscapes. In the same way, our smart soundscape system will engage the visitors by generating soundscape. However, the soundscapes composed for this thesis are based on the social gathering melodies instead of natural sounds. Further on, Astrid, et al. [4] is playing distinct soundscapes at different times of a year to study the new knowledge and scope of soundscape in various settings. On the other hand, this work is not based on any time constraints; the melodies will be played by mapping the visitor’s behaviors and interactions with the system.

Moreover, several other related works [8, 9, 11, 12, 13, and 16] describe multimodal system and interaction design contribution in museum perspective. However, they do not have relational soundscape for interactive environment nor any design scheme for generating the soundscape. Details of their contribution and their work intend in the field of HCI with respect to multimodal and interaction design.

2.3 Theoretical Framework

The conceptual design used in this thesis is based on an interactive environment scheme. This scheme focuses on why and how the visitors interact with one another and the system. It can be applied to extract input parameters and functional requirements for the system. As it consists of coordinated formation of multimodal resources, such as, nonverbal and paraverbal. Nonverbal resources include language, gesture, images, body movements; whereas, paraverbal contains prosody such as rhythm and sounds. These resources can be manipulated through the system to carry out interactional activities of the visitors. It enables the creation of an interactive space by the collaboration of the visitors with each other and with the system.

Furthermore, it explains different scenarios to play some sounds with corresponding parameters for the system such as movement, number of participants, and their distance from the artifact. It discusses the relationships among sounds along with the description of several situations, for instance, handling the timing of sounds to be played and stopped. It includes different semiotic modes of communication, such as, a sketch of different circumstances, short video clips of interactions, and linguistic context; therefore, it is named as Multimodal Design Brief (MDB).

The perspective of multimodal interaction is important for this work. The proposed theoretical framework includes the philosophy of multimodality and the multimodal systems. It gives insights into an idea behind multimodal designs. In the beginning, the concept of the MDB was not comprehensible for the author; however, the theoretical framework provided a foundation for an understanding of this idea. The term multimodality will be discussed in the following section.

2.4 Multimodality

A multimodal design has several modalities and complicated network of multimodal interactions. The word multimodality refers to various modalities which represent concepts like voice or sound, gesture, touch, vision etc. In other words, multi refers to more than one and modal indicates the notation of modality in the form of mode. It specifies the communication channels that help to deliver or to obtain information. It also refers to a kind of behavior to act, recognize or articulate an idea. Modes are beneficial to establish a way to understand or deliver the meaning of information [21]. An essential step in a multimodal design is relationship between information and modalities.

Furthermore, multimodal channels communicate their message using more than one semiotic mode or channels of communication. Films are multimodal because they use words, videos, music, sound effects, and moving images etc., even though all these are separate semiotic modes, however, together they represent one combined message. Moreover newspapers, magazines, and websites can also be defined as multimodal since newspapers include words, images and colors; magazines use words and pictures; websites use videos, sound clips, text, colors, pictures, and fonts [27]. Similarly, the MDB presented in this thesis is multimodal as it comprises various communication modes such as audio sounds, short films of visitor's interactions, and linguistic context of different events. These semiotic modes are connected with each other to form a final meaning.

Multimodal human-computer interaction (HCI) has become vital in many research areas including computer vision, psychology, artificial intelligence and much more. It's been acting as a connection between these research areas since computers have been integrated into several everyday objects (Ubiquitous and pervasive computing) [1]. In recent years, the multimodal gesture interfaces are relatively famous and playing a vital role in multimodal human-computer interaction. Multimodal computer systems support communication that allows users to communicate with different modalities such as voice, typing, and gesture. The author finds similar phenomenon in the smart soundscape system where the input is gesture based and the output is sound.

Multimodal interaction has mainly two views. The first view focuses on the human side that contains multimodal perception and control. The other view concentrates on the system side. On the human side, the term modality relates to both human input and output channels. The system side refers to constructing a based on the synergy of two or more computer input or output modalities. The process that translates sensory information into high-level illustration is called perception [21].

Humans use different interfaces, e.g. speech, aural, touch and mime (as appropriate to the situation) to communicate with their environments and with other people. Pickering, Burnhamt, & Richardson [5] suggest that combining these input modalities make the use of gesture more meaningful. For example, a voice with a gesture (a body movement) complements the meaning of

understanding the situation. Although there is not much multimodal motivated research available to authenticate this suggestion, this idea has been evidenced in the field of HCI [5].

Moreover, the body movements of the people in the environment can help to acquire the information regarding emotions and intentions. For example, a gesture is a linguistic form of movements, which are not using any written language or verbal articulation in a communication process. The gestures can be used to build up interactions between humans and systems [2]. In multimodal theory, it has been noticed that there are resources (e.g., language, gesture, images), which can lead to some understanding. In a multimodal message, the meanings are multiple articulates as it involves two (or more) resources such as images (static, moving), texts, sounds, and layout or design. Together they become a meaningful and functional as whole. Nowadays, most of the communication is multimodal and visual. The communication concept in the multimodal theory is therefore important. In next section, the author further will discuss the relationship between MDB and multimodal theory.

2.5 Communication in Multimodal Theory

Gunther Kress and Van Leeuwen (2001) have sketched four domains of practice called Strata. The strata comprise of discourse, design, production, and distribution [14]. In this thesis, the strata have an important role for determining meaning through analysis of modes and resources used in MDB. The author will discuss discourse, design, and production that are relevant to this work.

2.5.1 Discourse

According to Kress & Van Leeuwen (2001), discourse is socially constructed knowledge of some aspects of the reality. The knowledge of events is composed of the certainty that tells insight of social identities, for example, who, what, which, where and when etc. In the last two decades, a lot of research and analysis has been carried out on the discourse which resulted in at least two assumptions, discourse exists in language and that discourse exists regardless of any material realization [14]. Further on, James Gee (1990) writes in his book that the discourse can be used to classify a system language, such as, traffic regulation, or medical systems etc., [17]. There are several discourses interlaced in this work, for instance, the Swedish History Museum constructs a discourse. Similarly, the MDB articulates several events for the smart soundscape system.

In the software engineering, discipline and requirements are written in technical language to understand the functional aspects of a system. However, the MDB was not written into a system language discourse for developing the smart soundscape system. It contains several events and each event explains the sequence to play sounds and their relationships with each other. For example, a sound will be played when a visitor is standing inside a defined area in the museum and if the visitor changes his/her interaction, a new sound will be played accordingly. The author used an approach for discourse analysis to find out the meanings of a multimodal text. This approach is known as the overall discourse organization of a text. It immerses the techniques in which sentences are organized, for instance, the scenes and events make up the arguments for a given situation [17]. This approach helped the author to eliminate the ambiguities, complexities

and eventually translate the MDB into computer science discourse. The Swedish History Museum also constructs a discourse and it is known as the discourse of the museum. It provides information about designed space parameters. For example, visitors' interactions, visitors' locations and the museum artifact.

2.5.2 Design

According to Kress & Van Leeuwen (2001) said in their book, designs are the conceptualization form of semiotic products and events. In other words, a design is the abstract realization of discourse. It could be explained as the separate form of the actual material of the semiotic product or event. For example, the design is the layout of what is to be produced [14]. The MDB contains a composition in the form of a sketch which explains different events comprised in it. The author has stretched those events and labeled them with descriptive language, which helped to establish understanding about what is needed to be produced in the implementation of the system and which sounds will be played corresponding to each event. The third relevant domain of 'Strata' will be demonstrated in section 2.5.3.

2.5.3 Production

According to Kress & Van Leeuwen (2001), meaning (of multimodality) does not only reside in discourse and design but also in production [14]. The production is the communicative utilization of medium and material resources. It leads to actual instances of meanings by using these resources. Sometimes, it is hard to separate design and production, for example, while rehearsing for musical concerts, it is hard to know when performance begins and when design ends. However, the semiotic role can be often separated from production plays which are changeable in communications. For instance, a composer designs the music, but singers or performers get it done. The speakers or performers use body expressions to communicate directly while performing which adds a little meaning (that are not defined in the design by the composer) without losing essential information from design [14].

In this thesis, production involves in two levels; first level comprises the development of the smart soundscape system and output of the system. The second level depends on the human interaction with the system. The interaction between the system and visitors could be anything, for example, a gesture, voice, touch or any other action. The system would then map that particular interaction to produce the output (in the form of consonant sound).

Further on, the environment in this thesis could be multi-sensual as visitors not only hear sounds generated by the system but could see the object in the museum as well. Therefore, this work represents two different situations of multimedia modes for the visitors. First is the case investigation and second is the soundscape generation. The first mode is active, such as, poses, interactions and gestures to generate sound in the real time. Whereas, the second mode is also active, where, visitor can generate soundscape by collaborating with each other, while, interacting with an artifact in the museum.

2.6 Research Gap

In the past, much work has been done in the field of multimodal interaction design. In the previous works, different approaches are concentrating on providing the smart environment for the museums to increase their immersive experiences and exhibitions. For this purpose, the author has evaluated different peer-reviewed articles and concluded that they are not explaining any detail about interaction scheme, the implementation of systems regarding programming language, tools, and technologies used for implementation. In this research, the author would develop a system that can be implemented in different public places to increase user interactions. The author would also describe in detail how multimodal interaction design is interpreted that is directly related to soundscape composition solution.

Chapter 3: Research Methodology

3.1 Introduction

This chapter focuses on describing the ways to collect information and develop the system for the selected topic. Research methodology is an approach to collect data and make it meaningful to the research objectives. This thesis focuses on the translation and implementation of the MDB into a smart soundscape system. The complete specification of requirements of this system was not given in beginning. Therefore, the initial challenge was how to gather and understand the system's needs which are most suitable. We have presented the research planning and general structure for outputs of this dissertation.

3.2 Iterative Design Process

There are various techniques for gathering software requirements in the field of software engineering. However, we have applied new ideas, modifications, and improvements iteratively. This chapter describes how this thesis is organized by using the iterative design process. An iterative design process is chosen to conduct this work and it helped to acquire the required results. The main reason to prefer this method was to carry out this work into small portions and then review it to address the problems for further requirements. Additionally, the iterative nature of design process makes the ideas more suited for the appropriate solution. Since this approach is cyclic, so if a problem occurs during iteration, it can be fixed in next iteration by the improvements. To ensure that the problem is solved it is tested again. For instance, if a prototype is developed in the development phase and it is recognized that the conceptual design is not correct, it would be then fixed in a new iteration. In the same way, during the requirements analysis phase, if the requirements do not fully represent the problem, it is possible to go back to requirements gathering phase to solve the problem [7].

As shown in Fig. 1, four phases are combined with each other and a middle portion through arrows. The middle portion is named as assessment which represents the process of validating or evaluating each step. The author took the basic idea of Fig. 1 from [7] and modified the names of phases accordingly the needs of this work. Each phase has been explained in the following section.

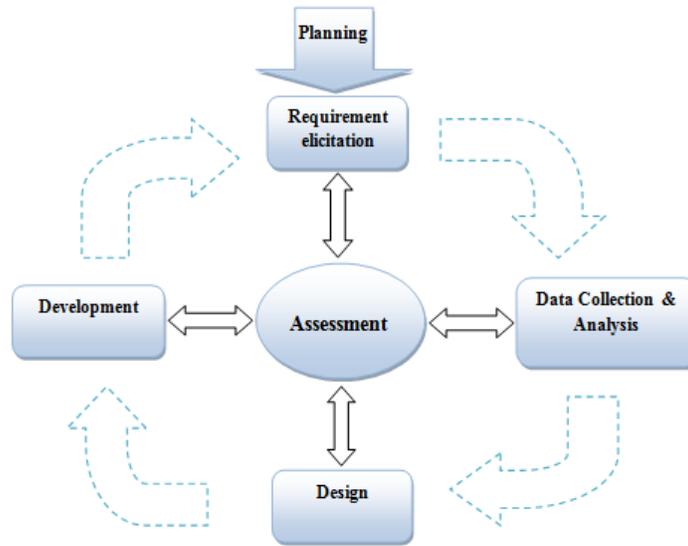


Figure 1: Research process Bill Wolfson, 2013

3.3 Data Collection (Primary and Secondary Data)

There are two types of the data known as primary and secondary one. Primary Data which is collected directly by researchers for a particular project is known as primary data [6]. In general, the primary data can be natural, observational or real obtained from questionnaires/interview. The data collected by the author in the thesis is primary and observational. On the other hand, secondary data is the data which is already collected by other researchers in this context, and the researcher analyzes that data to extract valuable information. In this research, the focus has been made on primary and the secondary research to develop and analyze the system.

This research was carried out with the collaboration of History Museum, Stockholm Sweden and it was decided to choose an artifact (See Fig. 2) of the museum as a trial object to collect the data. Therefore, the author collected the primary data by visiting the museum directly.



Figure 2: Stockholm history museum artifact chosen for project demonstration

The primary data consists of various interactions of visitors while looking the artifact in the museum. It helped to analyze the requirements for the system design. The author visited the museum during the summer for four to five hours to collect an adequate amount of data. The author observed twenty-five visitors and took notes of their observations. These observations are listed in Table 2 below.

For secondary data, several online databases of MDH University were used to search literature concerning this work. The sources which were used to collect relevant literature are (DiVA, Google Scholar, Scopus, ACM and IEEE Xplore). The author has found twenty-five different articles. After going through those articles, fifteen relevant articles have been chosen and reviewed. These articles were selected based on high quality research papers, empirical studies, and journals. Aside, some e-books and well-known websites were also explored.

The keywords used to search literature are embedded systems, multimodal, gesture technologies, human-computer interaction, soundscape, digital museum, audio applications, and interaction design. It was important to understand the concept behind multimodality in computer systems as the implementation of smart soundscape system is based on the MDB. A book “Multimodal Discourse” was helpful in building a strong base for the theoretical framework (See Chapter 2).

3.4 Sampling Approach

For sampling approach, the focus has been made on ten respondents. We observed these respondents with hidden eyes without letting them know. The reason is that, if people know they are being observed, they become extra conscious and this might impact on research’s results. The focus has been made on convenient sampling and only ten respondents were selected. It would be easy for the researcher to analyze them in less time. Certainly, the ethical considerations were taken into account while observing the visitors. The interactions observed by the author were not

recorded on any electronic media and random visitors were chosen. Their identities were also not used in any way.

Table 2 below shows the observations collected from the museum. These observations are necessary because the project required the movements and gestures of the visitors to be as natural as possible when interacting with the artifact. Therefore, we need to know what actions they perform when they are around the artifact.

Table 2: Observations collected from museum

Visitors	Observations
Gender: F Generation: Young	Two girls saw the stone and came close to it. One of them sat down and moved her right-hand fingers on red lines drawn on stone. Meanwhile she was also talking with her friend.
Gender: M Generation: Young	Two boys sat near the stone and they slightly pushed the stone one by one. Perhaps they were trying to know how heavy it is.
Sex: M Generation: Senior	An old man came and looked down to see the stone. His both hands were in his jeans pockets. After a while he moved to read description written about the stone.
Gender: M Generation: Young	When he was looking at stone, he started rubbing his shoulder and scratching his back bone.
Gender: F Generation: Middle-aged	When she was reading the description of stone, she suddenly put her left hand on her waist and slightly lowered her neck.
Gender: M Generation: Middle-aged	He was shaking his shoulder when he was looking at stone; perhaps he was enjoying museum's background music.
Sex: F Generation: Young	She was looking very curiously at the design carved on stone.
Gender: F Generation: Middle-aged	She came near to see the stone and then she moved back about 3 or 4 steps to take picture of it. While she was taking picture, she positioned her legs a little down and set focus of camera by her hand.

3.5 Data Classification and Analysis

Classification and analysis of collected data was done during a discussion. It was mandatory as the output was needed in designing and implementation of the system. In the discussion, it was decided to conduct an internal workshop of ten to fifteen minutes among the author and researchers. In the beginning, three minutes were given to each participant, in which they could think, analyze and then write possible interactions that could occur while visiting a museum artifact. Each participant described their written outcome in a listed sequence. Afterward, to visualize the written behaviors, it was decided to perform them virtually. For example, a participant would read his/her notes and other participants would act accordingly. The objective of this activity was to make sure that requirements met properly. It is fair to say that internal workshop was a good experience and it worked very well. We were able to analyze and validate the collected data and succeeded in identifying possible interactions.

Designing an efficient and error-free algorithm is a challenging task. It requires a powerful intellect for taking a problem and drawing a solution. There are plenty of books which describe different techniques to design algorithms [29]. As discussed in chapter 2, the algorithm implemented in this thesis is based on an interactive environment scheme named as MDB. In a course at the Information Design Department of the Mälardalen University, arts students were designing different soundscapes based on people's interactions. They were also targeting the museum's stone (See Fig. 3) for their coursework. We were invited to attend the final presentations of the arts students. During their presentations, we were particularly fascinated by a group of artists who simulated the Viking stone concept into a situation of dance along with modern party music generated by the interactions of visitors. This group was asked to work under the supervision of researchers. Artists agreed to work voluntarily for designing and composition of soundscape for the thesis. They have also formulated the MDB which is translated and implemented in this work. So, the output of their work complements the smart soundscape system and these two works are related with each other.

Furthermore, our list of selected interactions was later iterated and revised with more interactions suggested by the artists. To establish the meaning of proposed body movements and facilitate the artists to perform those movements in motion, they used gesture drawing (See below in Fig. 3). Gesture drawing approach is useful for the artists to well understand the utilization of muscles and typical range of gesture in the human body joints. Fig. 3 below shows a gesture drawing of the interactions proposed by the artists.

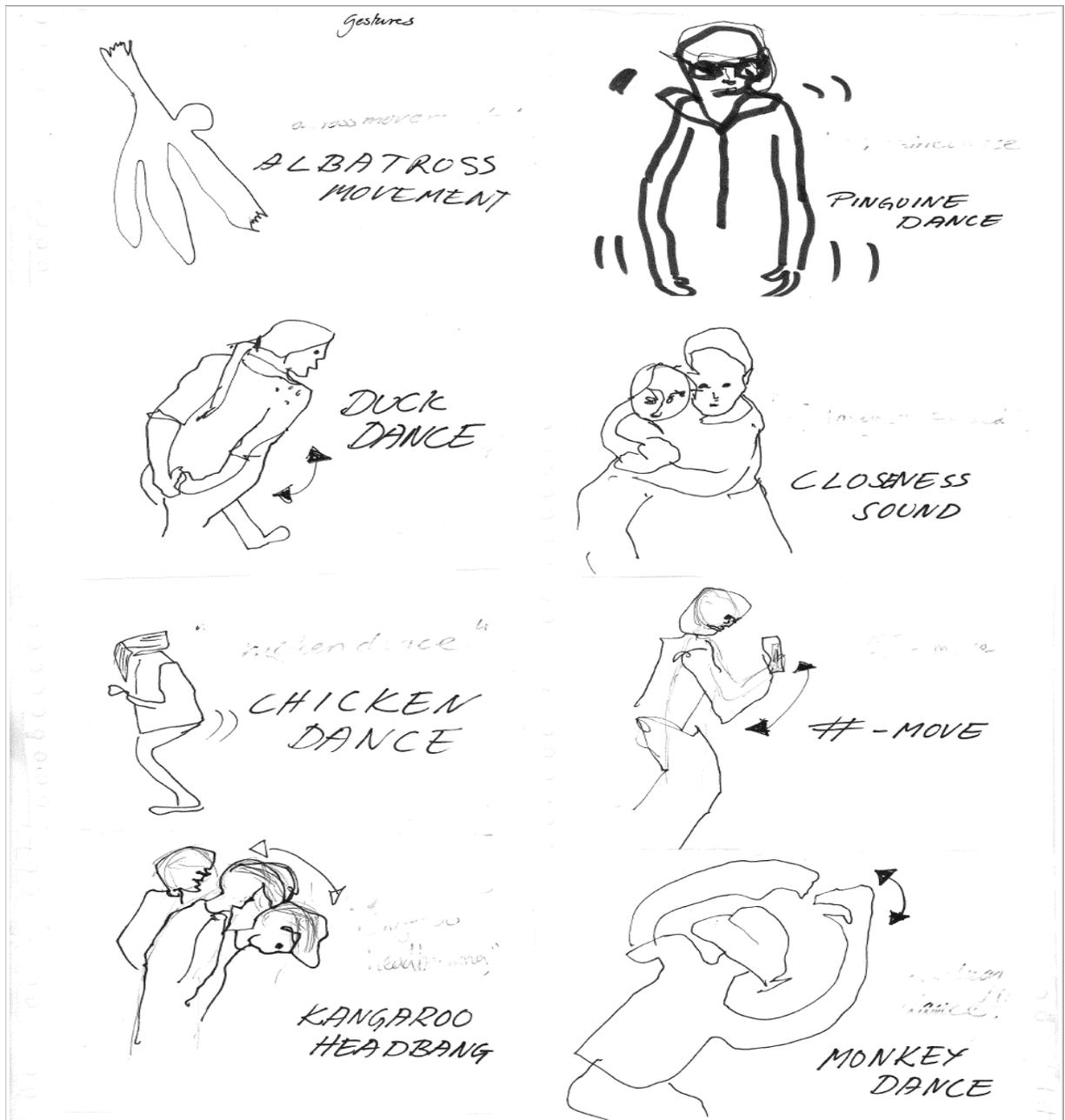


Figure 3: Gesture drawing to express different meanings of body movements

The proposed interactions were performed and recorded in short films. These short films are playing different sounds, i.e. background music, and these sounds are later used in this work. Artists used a cluster of nonverbal and paraverbal resources to provide an explanation to the interactions. Nonverbal resources are used for the creation of an interactive space through body movements and respond to this event, such as, the change in the environment. This exercise (See below in figure 4 and 5) played a key role in the establishment of the MDB.

Fig. 4 shows an example of a movement which indicate the course of interactions in response of different sounds, where (a,b,c) represent three different interactions (include walking, clapping, and Kangaroo head-bang) corresponding to each sound which is being played as background music.

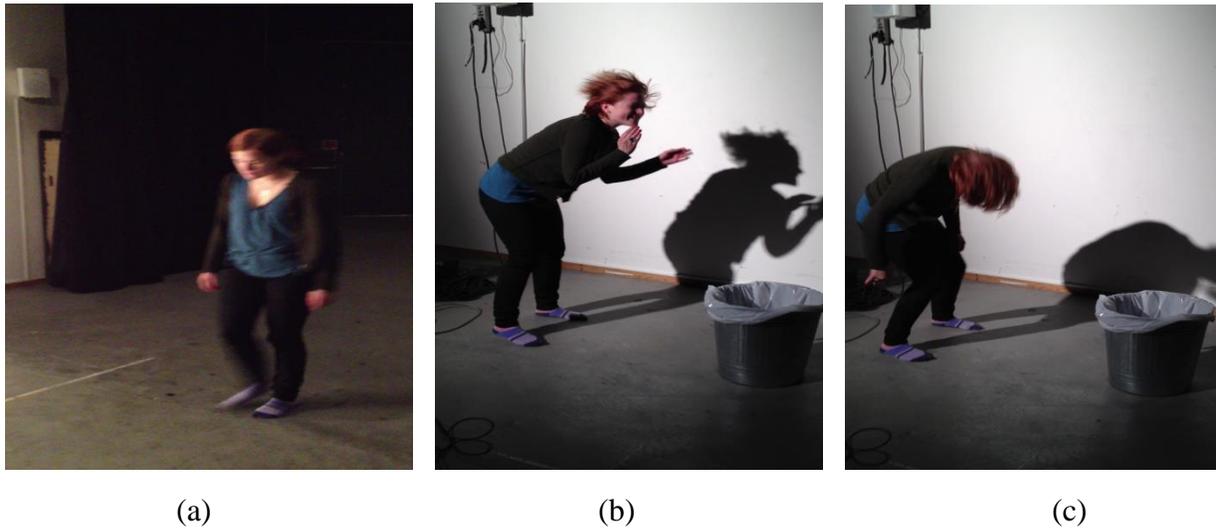


Figure 4: A segmentation of a sample interaction

Fig. 5 shows that an interactive environment which facility an artist to perform her arm and bod movements compatible with background sounds. where (x) represents duck dance and (y,z) a division of Albatross gestures, these are aforementioned in Fig. 3.

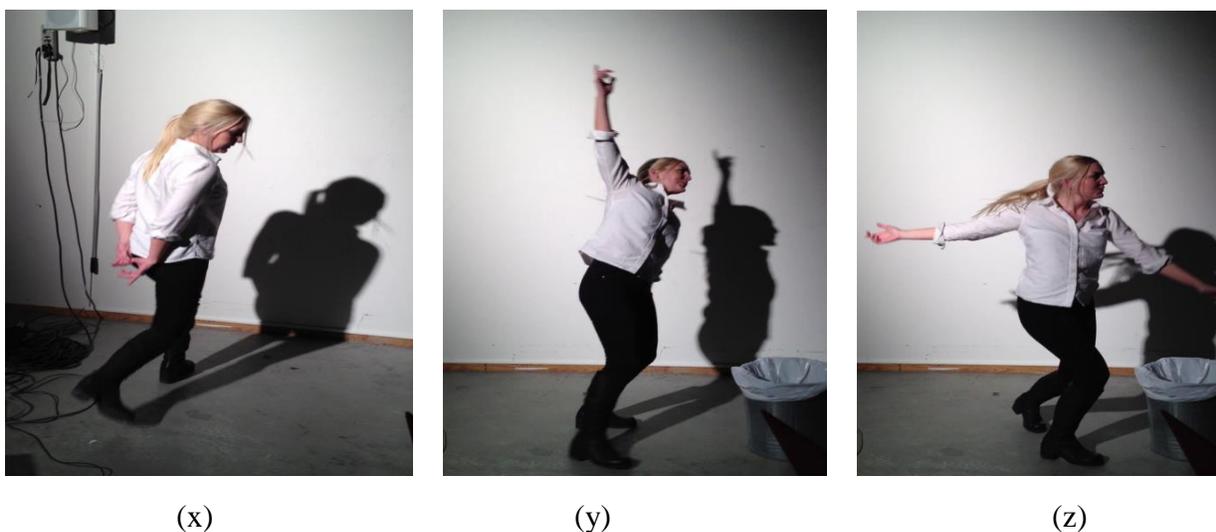


Figure 5: A segmentation of Albatross and Duck dance gesture

A list of interactions was selected from the received data in museum and the internal workshop (See Table 3). Table 3 below also contains artists' suggested interactions. Their suggested interactions are listed in bold and italics format to differentiate them from our findings, which established after the internal workshop.

Table 3: A list of interactions, selected by us from the collected data and suggested by students.

No	Interactions
1	Air drum
2	Approach of a grown up with a child
3	Approval head gesture
4	<i>Clapping</i>
5	Covering face
6	<i>Head banging</i>
7	Holding hands
8	Hugging
9	Jumping
10	Kicking
11	Pointing
12	Reaching hand
13	<i>Standing</i>
14	Surprise movement
15	Talking
16	<i>Taking photo</i>
17	Waving
18	<i>Walk around</i>

19	Walking away
20	<i>Head and hand banging</i>
21	<i>Beating hands in air</i>
22	<i>Swing hands left and right</i>
23	<i>Rear swing dance</i>

3.6 Design and Development

Design and development remain a challenging task for smart soundscape system. The challenge is due to get multiple input parameters such as number of visitors, their walking speed or interactions with the system and generate tailored sound in response. These parameters can affect the generated soundscape; for example, the system can play a sad sound with slow tempo on a walking speed of the visitor. Similarly, a head-banging interaction can allow the system to play fast sound etc. Regarding system inputs, we can use Kinect system as the problem could be solved best by it. The Kinect system can also help to track multiple visitors' interactions and recognize their interactions with the system. However, integration between Kinect and smart soundscape system remain a tough task. Regarding generation of tailored sounds, we can compose different sounds corresponding to our list of interactions which are listed in table 3 and mentioned above in section 3.5. However, we must develop a configuration interface where one can associate these compose sounds with possible inputs such as the number of people, their distance to artifact and interaction with the system.

In order to cope with these challenges, Microsoft (MS) Windows operating system (OS) has been chosen. To the best of our knowledge, this OS is most user-friendly for average computer user, several software or hardware work with a minimum of hassle and has support to run dot NET (.NET) framework. .NET framework is software framework and it is developed by Microsoft. This framework has been proposed as it includes a huge class library called Framework Class Library (FCL) and provides modern programming languages such as C Sharp (C#), VB.Net, C++ and J sharp (F#). FCL supports user interface, applications development (including web and desktop), database connective and data access etc. Microsoft proposed this framework to use for many new applications developed for MS windows platform.

As we can see, the Kinect system is best candidate for providing system's inputs and track multiple visitors' interactions. We want to integrate the Kinect system with smart soundscape system by installing it in the museum nearby a trial artifact. Kinect system is developed using FCL on MS Kinect Software Development Kit (SDK), which is officially developed by Microsoft. Therefore, having same operating system and technologies for smart soundscape and Kinect system will

eventually make the integration simple. Having same platform and development framework will also help to perform fast communication and reduce the chances of incompatibilities issues between both systems.

To this end, C# used as the primary programming language for the development of smart soundscape system. The target programming language (C#) has been proposed for several reasons, including integration with Kinect system, interactive debugging of the system and ease of user-friendly design. The C# programming language has different strengths and can incorporate with third-party libraries. In this work, C# is used for graphical user interface (GUI) and core implementation of the system. MS also built an integrated development environment for .NET framework applications named Visual Studio. Visual Studio has built-in user-friendly components. These components represent the shape of Windows form components such as buttons, icons, windows, fields, text, and labels etc. In this manner, Visual Studio facilitates the developers to create interactive and user-friendly applications with less effort. Therefore, in the development of smart soundscape system Visual Studio was used for user-friendly interface and writing C# programming.

Although, .NET framework provides audio application programming interfaces (APIs) and guidelines for developers to develop audio applications. However, it does not provide much built-in audio related utilities for fast development and manipulate audio files in an advanced way. Compared to this, an open source .NET audio library named NAudio, contains dozens of audio related classes to facilitate fast development of audio related utilities in .NET framework. This library includes large numbers of advanced features. In this work, it was used for changing the pitch and tempo of sounds. The library was also used to perform operations such as a smooth fade-in and fade-out of sounds.

As discussed previously, we need to develop a configurable interface which allows to create relationships among pre-compose sounds and different scenarios described in the MDB. However, storing, retrieving and managing these relationships is another challenge. In order to achieve this objective, we decided to use MS Access for database management. However, the database scale is small in this work. MS Access database is suitable for small-scale applications and supported by C# programming language. The data source was designed to build the system dynamic and facilitate the artists to manage the data. Therefore, one can simply perform different operations on data such as save, update and delete.

3.7 Author's Role and Contribution

The author was the sole lead developer on smart soundscape project. Therefore, author handled all aspects of this system such as designing, developing system and managing database under the supervision of the thesis supervisors. The author played a key role during this thesis; however, we did not participate in the design of MDB and sounds composition.

3.8 Research Ethics

We have considered different ethical aspects in this research. We did not disclose the name and information about respondents and their privacies are kept confidential. The data collected from various sources have been properly referenced. There is no copyright violation and we did not represent anyone else's work. All information is presented properly and there is no data manipulation to get desired results.

Chapter 4: Results and Discussion

4.1 Introduction

The results and discussion of the thesis are presented in next sections. The results of the thesis are divided into two parts. The **first part** explains the breakdown of the MDB into the system functional specifications. The **second part** describes to what extent these specifications are fulfilled. It also highlights the system’s functionalities and these functionalities are analyzed through the graphical user interface of the system.

4.2 System Overview

This thesis aims to investigate an algorithm and develop a system to map visitor’s behavior and interaction to the changes in the generated soundscape. The body movements or gestures will be collected using a Kinect system. The Kinect system is part of the MiM project and would be integrated with smart soundscape system. This system is developed from the translation of the MDB into functional requirements.

Fig. 6 shows an abstract sketch of the MDB and pre-compose sounds list. Small circles which are shown in Fig. 6, represent design space parameters such as number of visitors, their interactions, their locations and the museum artifact. Digits wrote inside these circles represent assigned sounds to each visitor. For example, a circle which illustrate digit 2 represents a sound (Beat Kickdrum) has been assigned to a visitor. As we can see in Fig. 3, this sound is labeled with digit 2 in the sounds list. In addition, if the visitor changes his/her interaction in response, then a sound which is labeled with digit 2,5 should be assigned to that visitor. This change has been illustrated in Fig. 6 with a circle which represents digit 2,5.

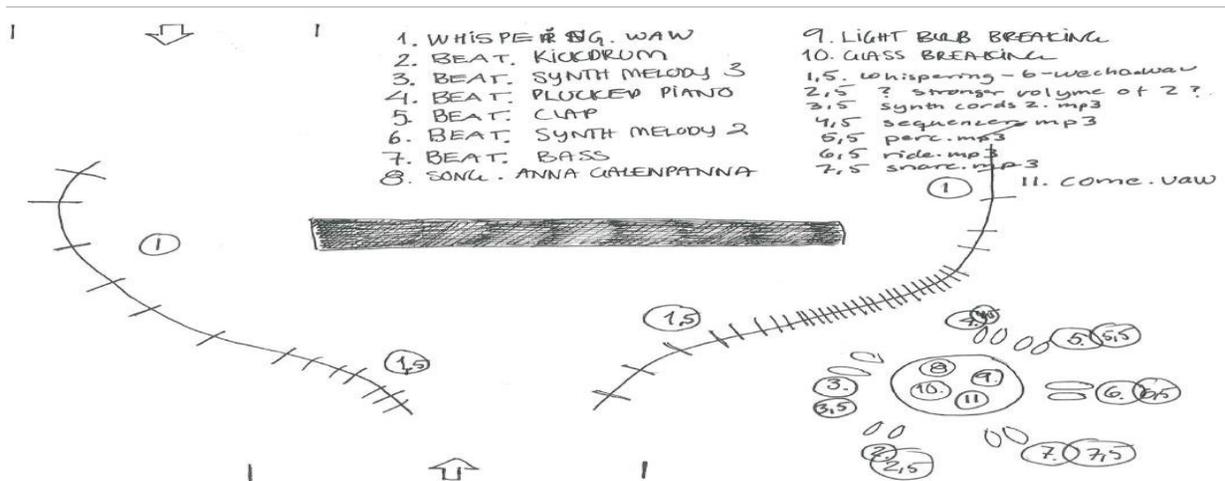


Figure 6: An abstract sketch of interactive environment

Initially as an experiment, it was planned to divide interactive space into six different spots in the museum around the trial artifact. And different sounds should be assigned to those visitors who are standing or changing interaction in defined spots. In this manner, visitors can collaborate with each other to create a soundscape together. However, maximum six sounds can be played together when there are six visitors in the defined boundary area. The MDB explains scenarios and different parameters that can best describe the visitor’s interaction and affect the generated soundscape. However, the MDB needs to be translated into a computer program. Therefore, the results of this thesis discussed below are the translation of the MDB into system’s requirements and a functional version of the smart soundscape system.

4.3 Translating Multimodal Design Brief (MDB)

This section summarizes the functional specifications of the system which are derived from the MDB. The MDB is based on several scenarios and consists of different semiotic modes of communication (See Section 3.2). The translation of these scenarios is listed in Table 4, 5, 6, and 7 and shown in Figure 4, 5, and 6 below. In each scenario, the best possible parameters are introduced such as the number of people, their interactions or distance to the trial artifact. They are used to generate a tailored soundscape corresponding to visitor’s interactions by associating them with composed sounds. The translation of the MDB is presented in tabular and graphical model. The graphical model helped to make the results concise, understandable, and verifiable.

4.3.1 Translating Scenario “Provoke by Standing Visitor”

In this scenario, the idea is that the system should generate a soundscape when at least one visitor is standing in a predefined spot. As discussed above (See Section 4.1), six different spots are planned to define in the museum around the trial artifact and different sounds are composed corresponding to each situation. These sounds can be played together to generate a soundscape. Each played sound should be assigned to each visitor (up to six) when they are standing in defined spots. The purpose of assigning sounds to the visitors is that system can change the soundscape which is further discussed in next scenario. Table 4 is describing the complete detail and translation of the present situation into functional requirement specifications.

Table 4: Sounds provoked by standing visitors

Sound names	Persons	Sound/s to be play	Description
Beat Kickdrum	1	kickdrum.mp3	The sound “kickdrum.mp3” should play and assign to a visitor when he/she is standing in defined spot near the artifact.

Beat Synth Melody 3	2	kickdrum.mp3 synthmelody3.mp3	The sounds “kickdrum.mp3 and synthmelody3.mp3” should play simultaneously when two persons are standing in defined spot near the artifact. Each sound should assign to each visitor.
Beat Plucked piano	3	kickdrum.mp3 synthmelody3.mp3 Plucked piano.mp3	The sounds “kickdrum.mp3, synthmelody3.mp3 and Plucked piano.mp3” should play together when three persons are standing in defined spot near the artifact. Each sound should assign to each visitor.
Beat Clap	4	kickdrum.mp3 synthmelody3.mp3 Plucked piano.mp3 clap.mp3	The sounds “kickdrum.mp3, synthmelody3.mp3, Plucked piano.mp3, and clap.mp3” should play simultaneously when four persons are standing in defined spot near the artifact. Each sound should assign to each visitor.
Beat Synth Melody 2	5	kickdrum.mp3 synthmelody3.mp3 Plucked piano.mp3 clap.mp3 synthmelody2.mp3	The sounds “kickdrum.mp3, synthmelody3.mp3, Plucked piano.mp3, clap.mp3, and synthmelody2.mp3” should play together when five persons are standing in defined spot near the artifact. Each sound should assign to each visitor.
Beat Bass	6	kickdrum.mp3 synthmelody3.mp3 Plucked piano.mp3 clap.mp3 synthmelody2.mp3 bass.mp3	The sounds “kickdrum.mp3, synthmelody3.mp3, Plucked piano.mp3, clap.mp3, synthmelody2.mp3, and bass.mp3” should play simultaneously when six persons are standing near the stone. Each sound should assign to each visitor.

Fig. 7 shows a graphical representation of a situation described in Table 4. As mentioned earlier, this scenario will be triggered when a visitor is standing in a defined spot. The system should play sound (kickdrum.mp3) then and assign this sound to that visitor. If another visitor enters the area afterward, the system should generate a new soundscape by playing sounds {kickdrum.mps, synthmelody3.mp3} together. These sounds should also be assigned to both visitors individually. Similarly, when three visitors are standing in the defined area, {kickdrum.mps, synthmelody3.mp3, plucked piano.mp3} should be played simultaneously and so on (up to six persons).

Fig. 7 describes a list of sounds and their combination with each other. The system will play each set of sounds together to generate different soundscapes. However, Fig 7 is providing an exemplary playing order of the sounds. The playing order of the sounds can be different depending on the system’s parameters and settings.

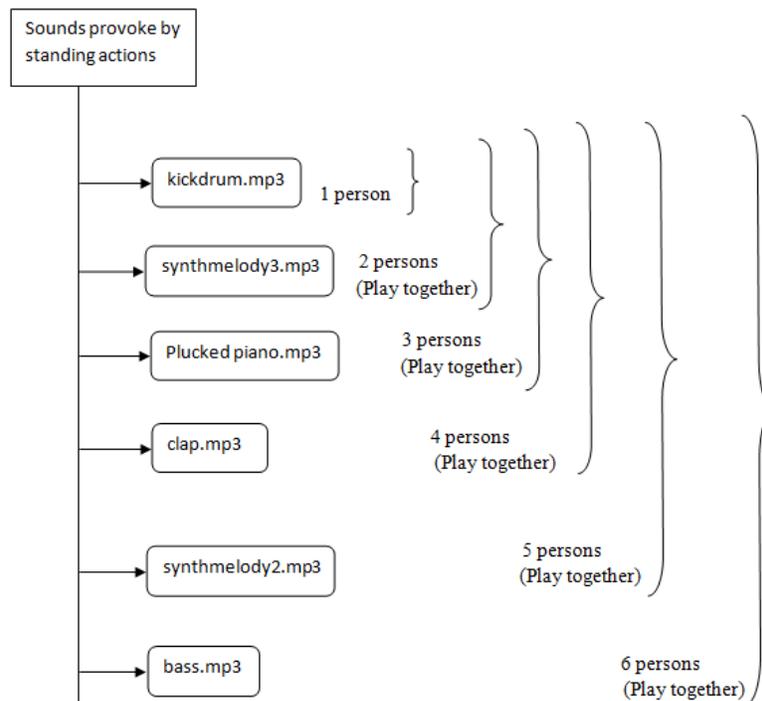


Figure 7: Relations between sounds provoked by standing visitor

4.3.2 Translating Scenario “Change in Interaction”

This scenario is an extension of previous situation explained above in Section 4.1.1 and both are interrelated with each other. The idea behind this scenario is that the system should generate a new soundscape when already standing visitors change their interactions in the response of

previously playing sounds. In other words, the change in their interactions should reflect the changes in the generated soundscape. The translation of this scenario is covered in more details in Table 5 below.

Table 5: Change in movements and change in sounds

Sound names	Sound already played or assigned to a visitor	Sounds to be played in response	Description
Beat Kickdrum	kickdrum.mp3	kickdrum.mp3	The system should play the sound “kickdrum.mp3” with high volume when 1 st standing visitor changes his/her interaction. This new change will then be played simultaneously with other sounds already being played.
Beat Synth Melody 3	kickdrum.mp3 synthmelody3.mp3	synthchords2.mp3	The system should stop already playing sound “synthmelody3.mp3” and a new sound “synthchords2.mp3” should play when 2 nd standing visitor changes his/her interaction. The new soundscape shall then be generated in accordance with already played sounds.
Beat Plucked piano	kickdrum.mp3 synthmelody3.mp3 Plucked piano.mp3	sequencer.mp3	The system should stop already playing sound “Plucked piano.mp3” and a new sound “sequencer.mp3” should play when 3 rd previously standing visitor changes his/her interaction in response of previously assigned sound. The change in his/her interaction should reflect the already generated soundscape and a new soundscape shall then be generated.
Beat Clap	kickdrum.mp3 synthmelody3.mp3 Plucked piano.mp3	perc.mp3	The system should stop already playing sound “clap.mp3” and a new sound “perc.mp3” should play when 4 th previously standing visitor changes his/her interaction (e.g. clapped hands) in response of previously assigned sound. The new

	clap.mp3		soundscape shall then be generated in accordance with already played sounds.
Beat Synth Melody 2	kickdrum.mp3 synthmelody3. mp3 Plucked piano.mp3 clap.mp3 synthmelody2. mp3	ride.mp3	The system should stop already playing sound “synthmelody2.mp3” and a new sound “ride.mp3” should play when 5 th previously standing visitor changes his/her interaction in response of previously assigned sound. The change in his/her interaction should reflect the already generated soundscape and a new soundscape shall then be generated.
Beat Bass	kickdrum.mp3 synthmelody3. mp3 Plucked piano.mp3 clap.mp3 synthmelody2. mp3 bass.mp3	snare.mp3	The system should stop already playing sound “bass.mp3” and a new sound “snare.mp3” should play when 6 th previously standing visitor changes his/her interaction in response of previously assigned sound. The change in his/her interaction should reflect the already generated soundscape and a new soundscape shall then be generated.

Based on the scenario discussed in Table 5, Fig. 8 shows the already playing sounds (which are mentioned in Fig. 7) and the corresponding sounds which will play by the system as soon as any of the visitor changes his/her interaction. For example, the sound “synthchords2.mp3” is assigned to a visitor and a new sound “synthmelody3.mp3” will play when that visitor changes his/her interaction. Similarly, if two visitors change their interactions in the same time then two sounds will be played, and previous sounds would be stopped (See Fig. 8). The braces in Fig. 8 are used to show which set of sounds will play to generate soundscapes based on visitors’ presence on the museum floor.

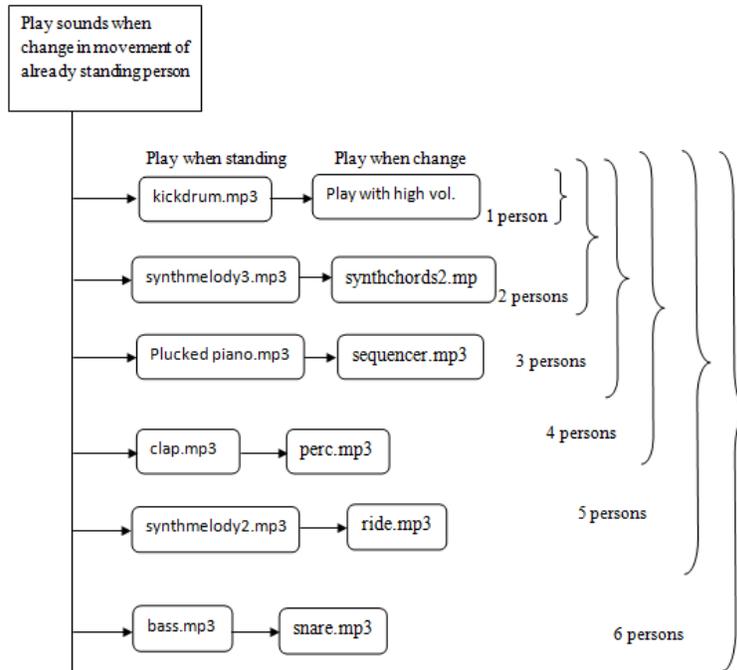


Figure 8: The change in sound by changing in body position

4.3.3 Translating Scenario “More than Six Visitors”

As discussed previously, the maximum sounds which can be played together are six when there are six visitors in the defined boundary area. However, what if there are more than six visitors in interactive space. In order to deal with this situation, a distinct scenario is introduced. Regarding this scenario, when there are more than six visitors gather near the trial artifact, the system will be allowed to generate a different soundscape instead of tailored. Thus, by applying this scenario all other sounds which are already being played should stop and entirely different sounds should be played one by one. The system’s functional requirements that are obtained from this scenario are listed in Table 6 below.

Table 6: More than six-person scenario

Sound names	Sounds to be play	Description
Anna Galenpanna	heavy_stone_slab_scape.mp3	The sound “heavy_stone_slab_scape.mp3” should play only when there are more than six persons near the trial artifact and all other sounds should stop immediately.
Light Bulb Breaking	Light_Bulb_Breaking.mp3	The sound “Light_Bulb_Breaking.mp3” should play when there are more than six persons near the trial artifact

		and the sound “heavy_stone_slab_scrape.mp3” completely played.
Glass Break Stephan	Glass_Break-stephan.mp3	The sound “Glass_Break-stephan.mp3” should play when there are more than 6 persons near the trial artifact and the sound “Light_Bulb_Breaking.mp3” completely played.

The graphical representation of Table 6 is shown in Fig. 9. There are six different sounds that should be stopped as represented in Fig. 9. This figure also shows the sequence of other sounds which should be played one by one when there more are than 6 persons near the trial artifact. Fig. 9 shows an example model of this scenario however; the sounds can be played in different sequence according to the system settings and artist’s plan.

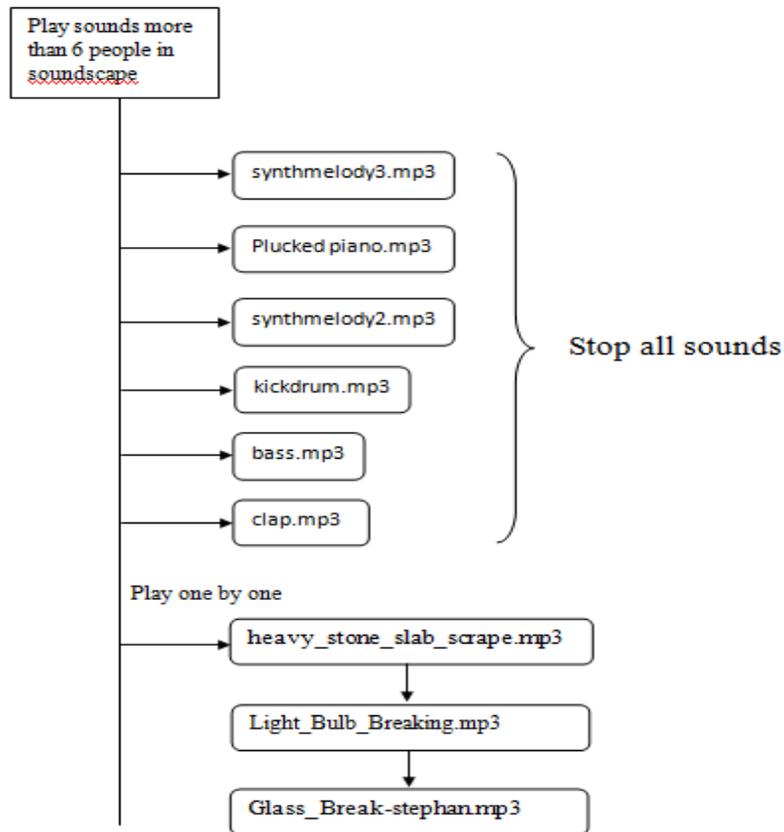


Figure 9: Play sounds for more than 6 people in soundscape

4.3.4 Translating Scenario “Based on Distance”

The final scenario of the MDB is slightly different and based on the visitor’s distance from the trial artifact. In addition, it also deals with other possibilities for example, what if all presented visitors go away from the defined area in the museum. The translated functional requirements of this scenario are listed in Table 7 below.

Table 7: Other sounds

Sound names	Sounds to be play	Description
Whispering	four-voices-whispering.wav	The sound “four-voices-whispering.wav” should play when there is 10-meter distance between visitor and trial artifact.
Whispering Echo	voices-whispering-6-wecho.wav	The sound “voices-whispering-6-wecho.wav” should play when a visitor coming close to the artifact and distance would be about 2 to 3 meters between them.
Come	come.wav	The sound “come.wav” should play when everybody goes away from defined area around the artifact.

4.4 Functional Requirements Implementation and Graphical User Interface

The smart soundscape system implementation has been achieved by implementing technical requirements which are translated from the MDB. The system is a desktop application and developed for Microsoft Windows. In order to focus on developing the system’s backend logic and user interface graphical (GUI), Microsoft .NET technologies were used (See Chapter 3). The main purpose of the GUI is to allow the artist to manage the configurations of the system and try out different combinations and scenarios (See Fig. 10).

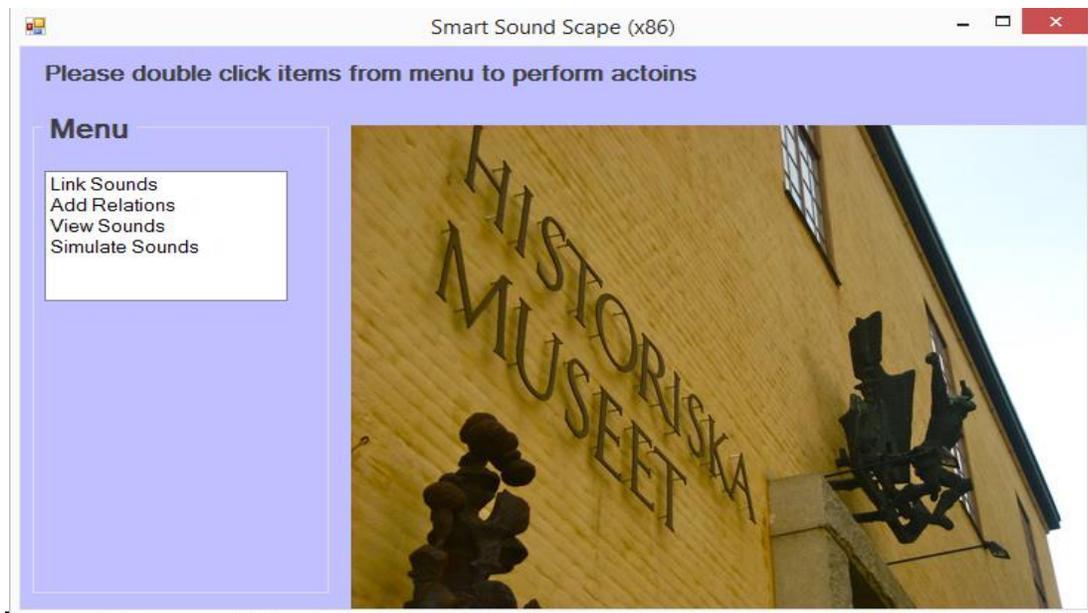


Figure 10: System's main menu screen

Link Sounds

This configuration facilitates to associate pre-composed sounds with different parameters such as the number of people, their distance to artifact and interaction with the system (See Fig. 11). The smart soundscape system is generating a soundscape based on these parameters. By connecting them with each sound, the complexity of the system has been simplified. It would be possible to make the system functional according to the proposed MDB through this approach.

Further on, the smart soundscape system will be integrated with the Kinect system. Therefore, the sounds' association will help the system to create a tailored soundscape for the environment by taking the input from Kinect system. The Kinect system has ability to obtain the information from the surroundings with the assistance of sensors. This information can be the number of visitors, their distance or interactions with the system.

Fig. 11 shows that the Link Sounds configuration also allows to setup other properties for example, playing order of sounds and conjunction. The purpose of such properties is that the system can run scenarios of the MDB which do not require association (See Section 4.1.3 and 4.1.4).



Figure 11: Screen which facilitates to link sounds

Add Relations

As discussed in section 4.1.2 above, the change in the visitor interaction should affect the generated soundscape. For this purpose, it is important for the system to analyze which sound/s should be played. To make this possible the application allows for defining the relation among different sounds (See Fig. 12 below). Through this approach we can define the relationship between two sounds by selecting parent and child sounds. The parent sound is the main sound for a scenario and the child sound would be played when a change occurs. However, it is the Kinect system responsibility to notify the smart soundscape system about the change in the visitors' interactions.



Figure 12: For adding/setting up relations

Simulate Sounds

Testing the soundscape setup easier remains a challenging task for the artists. The challenge is because, the smart soundscape system and Kinect system are not integrated. To deal with this challenge, we implemented a simulator to perform the tests (See Fig. 13). In order to validate all the translated functional requirements proposed in the MDB (Section 4.1.1, Section 4.1.2, Section 4.1.3, and Section 4.1.4), some tests were designed. These tests cover all types of scenarios presented in the MDB. The following tests are conducted based on functionality testing approach and verified that the system is producing desired results.

Test Case-1: Provoke by standing – Soundscape generating through standing interaction.

Test Case-2: Change in interaction – Affect the generated soundscape in the response of visitor’s new interaction with the system.

Test Case-3: Provoke by many people - Play different sounds at the same time when more than six visitors are present on floor.

Test Case-4: Distance based – Change the generated soundscape when visitors come close to the artifact within ten-meter distance range.

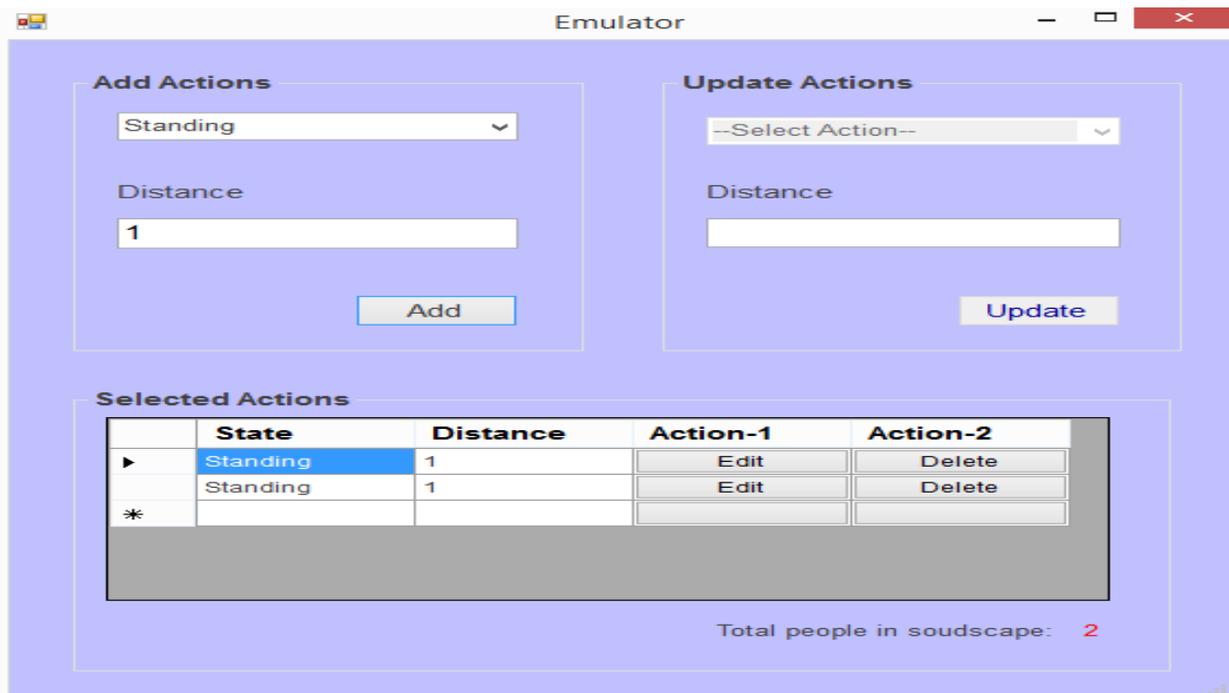
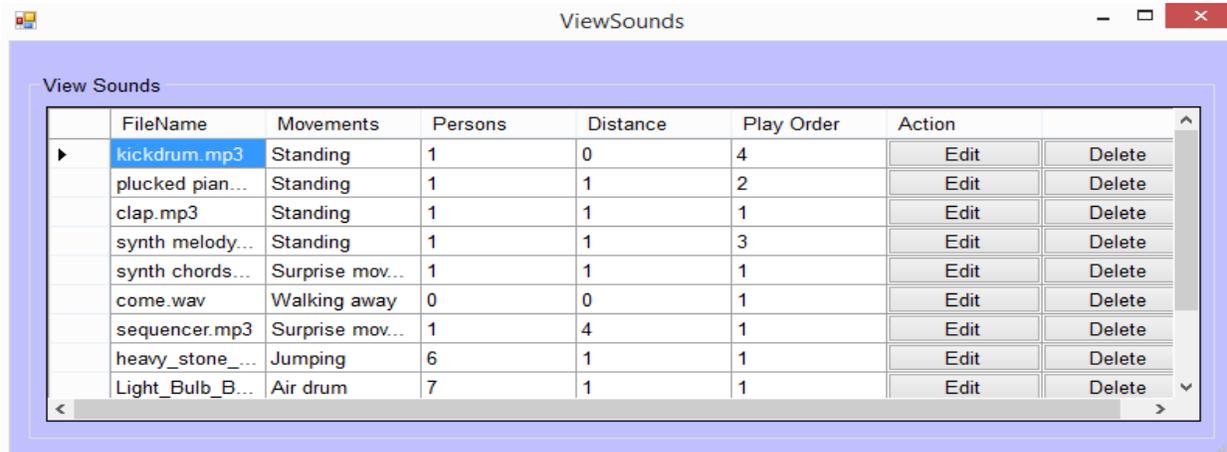


Figure 13: Simulator for testing generated soundscape

View Sounds

The view sounds option loads a list of all sounds stored in the system corresponding to their association with different parameters (See Fig. 14). It is a general functionality which allows the user to view, update or delete the sounds' associations.



The screenshot shows a Windows application window titled "ViewSounds". Inside the window, there is a table with the following columns: FileName, Movements, Persons, Distance, Play Order, and Action. The Action column contains "Edit" and "Delete" buttons for each row. The first row is selected, and its "FileName" is "kickdrum.mp3".

FileName	Movements	Persons	Distance	Play Order	Action
kickdrum.mp3	Standing	1	0	4	Edit Delete
plucked pian...	Standing	1	1	2	Edit Delete
clap.mp3	Standing	1	1	1	Edit Delete
synth melody...	Standing	1	1	3	Edit Delete
synth chords...	Surprise mov...	1	1	1	Edit Delete
come.wav	Walking away	0	0	1	Edit Delete
sequencer.mp3	Surprise mov...	1	4	1	Edit Delete
heavy_stone_...	Jumping	6	1	1	Edit Delete
Light_Bulb_B...	Air drum	7	1	1	Edit Delete

Figure 14: View sounds relationships

4.5 Discussion

The findings of the thesis in relation to the theoretical framework and previous research are discussed in this section. The main part of the discussion is based on methodology which is used to accomplish this thesis and eventually produce some meaningful results. The expected goals of the thesis have been achieved by the development of smart soundscape system. And results show that within the scope of the system all requirements are implemented. Moreover, a test simulator was implemented to conduct a series of tests and determine the system's output. The aim of providing this simulator is also to analyze the parameters which can best represent visitor's interactions and affect the generated soundscape. After having performed different tests (See section 4.2), we have been able to conclude that the system generates tailored sounds based on different parameters such as number of visitors, their walking speed or interactions and these parameters affected an interactive environment.

Numerous challenges were encountered in this thesis. Initial challenge was to understand the multimodality as the author had no prior knowledge of multimodal interaction design field. However, the literature review and theoretical framework have been of great importance in establishing author's basic understanding of this concept. Furthermore, in the development phase it was required to develop a Windows based desktop application and implement advanced features to manipulate the sounds effect for example, changing the sound tempo and perform fade in/out.

It obliges the author to learn how to implement such applications and features. In order to achieve this, the author explored different development forums on the relevant areas and they helped to get started with the development work.

As discussed in chapter 3, the methodology used in this thesis is an iterative design process. The rationale to choose this method is based on the facts for instance, it implements iteration which ensures the refinement of each process. In addition, it receipts the fulfillment of each individual task. This methodology helps to improve the processes and implementing innovative ideas to meet the correct requirements or the solution of a problem. In this manner, the author has divided this thesis into four different processes (See Fig. 1 in chapter 3) and performed several iterations to improve those processes. It is fair to say that after the completion of each iteration, there were visible refinements. The iterations were applied after getting feedback from the researchers and artists in the thesis. Thus, this method also provided a better way to get feedback from stakeholders.

Furthermore, optimization of software design improved by using the interactive design methodology. Since this approach involves iteration which means if there is an error introduced in one task it will not precede to the next task. Similarly, it is witnessed in this thesis that initial development of the system was insufficient to meet the degree of performance and quality required. As in the beginning of this work, the complete specification of requirements of this system was not specified. Therefore, the problem with this version was, it did not produce desire output corresponding to inputs. To this end, the iterative design method was an excellent approach to add innovative ideas and refinement iteratively throughout in this work. By dividing this method into separate phases, the main problem was solved by analyzing the correctness of each phase on every feature individually. This method not only helped us to solve the problem, but also played a vital role to achieve the expected goals.

Apart from the benefits of this methodology, the main challenge with this approach, it is mandatory to perform a solid and comprehensive analysis in the initial stages which would otherwise lead to costly process changes with minor or even no improvements [18]. The author finds similar challenges with the method in this study. The translation of the MDB to a computer program was a challenging task for us. The challenge was, due to several modalities, complicated network of interactions and inputs to the system. Initially, the author did not fully understand the descriptions and representations of the idea. It was revealed while presenting a prototype of the system to the artists. The reason for a lack of understanding was that the language and the representations to the design discourse were not linked directly with the system requirements. As the artists did not have computer science background, therefore it was decided that the artists will present their exhibition in design process and the author will translate their idea into software requirement specifications (SRS). In order to achieve this objective, we did a complete analysis of presented demonstration of their idea. Afterwards, we proposed a SRS which is the translation of the MDB. The assessment process was advantageous to verify SRS before to start development iteration. This SRS is used for the development of the system. It is fair to say that iterative design approach helped to improve

the quality of SRS by repeating development process until the required degree of quality and performance achieved.

As we can see, the smart soundscape system implementation was performed by considering the SRS using Microsoft technologies. The implementation of the system is divided into separate layers and it would be useful for code maintenance in future development or enhancements of the system.

Chapter 5: Conclusion

This thesis tries to investigate the possibility of developing a system that can interact with people near it, through composing and playing unique sounds for them. The research question, which gave directions to achieve the purpose of this thesis is as follows:

“How to break down an interactive environment scheme into functional requirements, so that they can be implemented for smart soundscape system?”

To answer the research question, we formulated some research objectives (See section 1.3). Therefore, the first section of this chapter describes overview of results of research objectives and its main contributions. In the latter section, future work that can be carried out for further research is presented.

One of the objective of this thesis is to investigate parameters that can best represent the visitor's interaction and they can affect the generated soundscape. To achieve this objective, the author collected the primary data by visiting the museum directly and observed visitor's interactions with a trial artifact. This work was extended through analyzing gathered data and later the data was revised with a group of artists. So collaboratively a list of possible visitors' interactions with an artifact are provided (See Table 3 in Section 3.5). During the data classification and analysis process, we conducted an internal workshop among us. The workshop was organized with two aims, the first aim was to analyze and validate the collected data. The latter aim was to find the best parameters which can be input to assist the system. So, based on our observations while visiting the museum and workshop findings we purposed some best parameters. And these parameters are visitors' interactions, number of visitors, and their distance from a trial artifact.

In Section 2.3 an interactive environment scheme (named as MDB) is mentioned, it is an idea and designed by a group of artists. This is accomplished by discussing the theoretical framework with artists. So, relevant theories were huge source of information for designing this scheme and it helped to achieve the objective; how to design an interactive multimodal scheme. This scheme comprises multimodal resources such as nonverbal and paraverbal and it discusses the relationship between these resources. Furthermore, the pre-composed sounds which are in initial version in the thesis, were also composed by that group of artists. Therefore, the output of artists work complements this work and these two are interrelated with each other. The composition of sounds has done through a proposed list of visitors' interactions (See Table 3 in Section 3.5).

In this thesis, the above mentioned an interactive multimodal environment scheme is a prerequisite for formulation of the functional requirements. One should have enough understanding about a multimodal design so that it can be precisely break down into a meaningful course. The knowledge is obtained by investigating the literature review about multimodal systems. Also, literature is reviewed to know the opinion of other researchers about the multimodal designs and how they defined multimodal environment. In this work, the MDB analyzed through given semiotics modes

such as sketches, videos, and textual data. These modes are important to break down this scheme into the functional requirements for the development of the smart soundscape system. Furthermore, to get insights behind this design and propose requirements that are applicable for the development of the system, we conducted meetings with artists, as it was very important to question the artists who were participated to design this scheme. During the design requirements process the author faced a challenge, what is the best way to describe complete, consistent, correct, and unambiguous design requirements. There is, though, no specific technique to write quality requirements [29]. We used a tabular approach to represent the MDB translation into software requirement specifications (SRS). Therefore, different situations explained in the MDB were break down as SRS in tabular form (See Table 4, 5, 6, and 7). This approach is convenient to classify functional requirements (i.e. MDB results in this work) when each requirement contains combination of parameters.

Moreover, we also draw a graphical model (See Chapter 4, Fig. 7, 8, and 9) to represent MDB's scenarios. The graphical model of each situation comprises of parameters such as visitor's interaction, tailored sound, and number of visitors. The graphical models helped to make the scenarios concise, understandable, and verifiable.

The MDB results were discussed with the artists. It was decided to implement MDB results for development of the smart soundscape system. Microsoft tools and technologies (include windows platform and .NET framework) played a huge role to develop the system. These technologies were chosen based on motivate design decision to make the integration easy of this system and Kinect system. One challenge has been experienced, how this system can simulate and generate the soundscape without integrating it with Kinect system. To address this issue, the developed system provides an independent interface to test the functionality of the system where one can apply different parameters such as number of people and their behaviors etc. This simulator helped to experience the system's efficiency and accuracy. Smart soundscape system is working according to MDB results. This is confirmed by providing a demo to the artists. Furthermore, an artist installed the system in museum and it was there for six months, though we directly were not involved. So, it can be concluded that developed system was used to generate the final soundscape for the museum.

Finally, we used a general iterative design approach to address the problems and achieved the research objectives of this thesis. This approach is simple, and it brings possibility to get feedback from stakeholders in each iteration to make sure that requirements are met. That way we not only identify their needs properly but also improve and generate innovative ideas to fulfill their needs.

5.1 Future Work and Suggestions

The research aim was to investigate and develop a smart soundscape system. The implementation of this system is based on a proposed MDB. However, the MDB has limited scope for generating soundscape based on multimodal interactions of the museum visitors with the system. Therefore, the future research can be carried out to address the limited scope of it by adding more situations

or scenarios. In future following features can be considered for the further development of the system:

- The current implementation of the system is for assigning sounds, up to six people and one sound for one person. Therefore, the relationship between sounds and people is one to one which could be enhanced to one-to-many relationship.
- In addition, more visitors' interactions could be defined corresponding to tailored sounds for the system.
- The current version of the system is taking care of changes that could be generated in soundscape up to two levels. In future it could be a good addition to implement more levels.
- Lastly, the author also suggests that Graphical user interface design has room for improvement in the future development of the system.

References

- [1] Alejandro Jaimes and Nicu Sebe, (2005), Multimodal Human Computer Interaction, *A Survey, IEEE International Workshop on Human Computer Interaction in conjunction with ICCV* on Oct. 2005, Beijing
- [2] Andrew W. et al (2006), ‘Choreographic buttons: promoting social interaction through human movement and clear affordances’, *MULTIMEDIA '06 Proceedings of the 14th annual ACM international conference on Multimedia*, 23 Oct. 2006
- [3] Andrew Sedgwick, Chris Luebke and Josef Hargrave, *Museums in the Digital Age*, published by: Arup Foresight + Research + Innovation (October 2013) [Online]
Available at:
https://www.arup.com/-/media/arup/files/publications/m/fri_museums_in_the_digital_age.pdf
(Accessed 11 June, 2016)
- [4] Astrid Drechsler, Hannes Raffaseder and Barbara Rubisch (2012), ‘Klang.Reise – New scientific approaches through an artistic soundscape installation?’, *a Conference on Interaction with Sound, AM 2012 - In Cooperation with ACM SIGCHI* on 26 September 2012, Greece: ACM
- [5] Carl A. Pickering, Keith J. Burnhamt, Michael J. Richardson (2007), ‘A Research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction’, *Automotive Electronics, 2007 3rd Institution of Engineering and Technology Conference* on 28-29 June 2007 Warwick, UK
- [6] Donald Currie, *Developing and Applying Study Skills*, published by: Chartered Institute of Personnel & Development (February 1, 2005) [Online] Available at:
<http://www.cipd.co.uk/NR/rdonlyres/E4D6775E-07B6-4BCF-A912-C3DE563C3F74/0/1843980649SC.pdf> (Accessed 23 June, 2013).
- [7] Bill Wolfson, *Engineering Design* Available at: http://www.integratingengineering.org/top_page.htm (20 Dec. 2013).
- [8] Eva Hornecker and Matthias Stifter, 2006, ‘Learning from interactive museum installations about interaction design for public settings’ *OZCHI '06: Proceedings of the 18th Australia*

conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments
on 20 Nov. 2006, Sydney, Australia

- [9] Eric, W. et al (2004), 'Augmenting the science centre and museum experience' GRAPHITE '04 *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, 15 June, 2004
- [10] Fakhreddine Karray et al. (2008), 'Human-Computer Interaction: Overview on State of the Art', *International Journal on Smart Sensing and Intelligent Systems*, March 2008
- [11] Gau, M. et al. (2012), 'Responsive Open Space Sound and Image in Public Spaces', *Virtual Systems and Multimedia (VSMM), 2012 18th International Conference on* 2-5 Sept. 2012, Milan Italy
- [12] Giovanna Varni, Gualtiero Volpe, and Antonio Camurri, 2010 'A System for Real-Time Multimodal Analysis of Nonverbal Affective Social Interaction in User-Centric Media', *Multimedia, IEEE Transactions on*, Volume: 12, Issue: 6, Oct. 2010
- [13] Gonçalves, Lígia, Pedro Campos, and Margarida Sousa, 2012, 'M-dimensions: a framework for evaluating and comparing interactive installations in museums', *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*. 14 Oct, 2012, Copenhagen, Denmark
- [14] Gunther Kress, Theo van Leeuwen (2001), *Multimodal Discourse*. London: Hodder Education
- [15] Hiromi T. Tanaka et al. (2010), 'Multimodal Digital Archiving and Reproduction of the World Cultural Heritage Gion Festival in Kyoto' *9th ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry* on 12 December 2010, South Korea
- [16] Ignore Aini Zolkifly, Lim Kok Cheng and Zasriati Azla Sabot, 2008, *Virtual Reality Musical Museum*. A conference on Computer Games and Allied Technology 08: Animation, Multimedia, IPTV and Edutainment on 28 April 2008, Singapore
- [17] James Paul Gee (1990), *Social Linguistics and Literacies: Ideology in Discourses* London: Falmer Press

- [18] Johnson, E.W., Brockman, J.B., Vigeland, R. (1996), Sensitivity analysis of iterative design processes, *Proceedings of the 1996 IEEE/ACM international conference on Computer-aided design*. 10-14 Nov. 1996, San Jose, CA, USA
- [19] Karen Johanne Kortbek, 2008, 'Interaction design for public spaces', *MM '08 Proceedings of the 16th ACM international conference on Multimedia* on 26 Oct, 2008, Ottawa, Canada.
- [20] Karen Johanne Kortbek and Kaj Grønbaek, 2008, 'Communicating art through interactive technology: new approaches for interaction design in art museums' *NordiCHI '08 Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* on 20 Oct. 2008, Lund, Sweden.
- [21] L. Schomaker et al. (1995), *Taxonomy of Multimodal Interaction in the Human Information Processing System*. A Report of the Espirit Basic Research Action 8579 MIAMI, February 1995
- [22] Lucia Terrenghi and Andreas Zimmermann, 2004, 'Tailored audio augmented environments for museums' *IUI '04 Proceedings of the 9th international conference on intelligent user interfaces* on 13 Jan. 2004, Brussels.
- [23] Matthew MacDonald (2012), *Beginning ASP.NET 4.5 in C#*. Apress. pp 8-16
- [24] MasIE, Museums as intelligent environments. Available at: <http://hci.ece.upatras.gr/masie> (06th Nov. 2014)
- [25] msdn.microsoft.com, *Microsoft Development Network* 2013. Available at: [msdn.microsoft.com/en-us/library/8bxxxy49h\(v=vs.110\).aspx](http://msdn.microsoft.com/en-us/library/8bxxxy49h(v=vs.110).aspx) (07th Dec. 2013).
- [26] NAudio Documentation Wiki, *CodePlex*. Available at: <http://naudio.codeplex.com/documentation> (14th Dec. 2013).
- [27] OpenLearning, Labspace. Available at: <http://labspace.open.ac.uk/mod/oucontent/view.php?id=445539§ion=4.1> (28 Feb, 2014).
- [28] Steven Conn (1998), *Museums and American Intellectual Life, 1876-1926*, the University of Chicago Press, 262

[29] Steven S. Skiena (1997), *The Algorithm Design Manual*, Department of Computer Science
State University of New York.

[30] Writing Quality Requirements, Karl E. Wiegers. Available at:
<http://www.processimpact.com/articles/qualreqs.html> (16 Jan 2016)