RWTH AACHEN UNIVERSITY

Movement analysis of visitors using location-aware guides in museums

Master's Thesis at the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Sebastian Borggrewe

Thesis advisor:
Prof. Dr. Jan Borchers

Second examiner:
Dr. Mark Wright

Registration date: 21.02.2013
Submission date: 28.03.2013

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed, and that I have marked any citations accordingly.

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, March 2013
Sebastian Borggrewe

Contents

List of Figures

List of Tables

Abstract

Understanding human behaviour has a long tradition in the field of psychology, sociology and also human-computer interaction. The behaviour that has been looked at in this thesis, is the movement of visitors in a museum, the historical city hall in Aachen. In particular I wanted to explore whether multimedia guides have an impact on human movement and behaviour in a museum.

I implemented a study to record visitors, using cameras in 6 rooms of the city hall in Aachen. I implemented two pieces of software, one to annotate the recordings, in order to extract the visitors movement path and one to analyse the movement data with the help of visitor's paths and heatmaps.

Based on these results I discovered that there are certain types of behaviour and also patterns that can be linked to the use of multimedia guides. However I also found behaviour that is rather linked to group visitors and individuals than the guides itself. This thesis covers related work, the study conducted, the software implemented, the evaluation and summary of our findings, as well as future work.

Überblick

Das Verstehen von menschlichem Verhalten hat eine lange Tradition in der Psychologie, Soziologie und auch in der Mensch-Computer Interaktion. Das Verhalten, welches diese Thesis behandelt, ist dass von Besuchern in einem Museum, des historischen Rathauses in Aachen. Ich wollte herausfinden, ob die Benutzung von Multimedia Guides einen Einfluss auf Bewegung und Verhalten in einem Museum haben.

In einer Studie wurden Besucher des Museums in 6 Räumen durch Kameras aufgezeichnet, um diese Videos mit Hilfe von Software zu analysieren. Es wurden zwei Applikation implementiert. Die Erste wurde zur Annotation der Videos genutzt, um den exakten Pfad der Benutzer im Museum zu extrahieren. Die Zweite dient der Analyse der resultierenden Bewegungsdaten und der Visualisierung durch Pfade und Heatmaps.

Basierend auf diesen Ergebnissen fande ich heraus, dass es gewisses Verhalten und Verhaltensmuster gibt, die dem Nutzen von Multimedia Guides zugeordnet werden können. Außerdem konnte ich Verhalten beobachten, welches Gruppen- und Einzelbesuchernspezifisch war. Diese Thesis behandelt den Stand der Forschung, die von mir durchgeführte Sudie, die implementierten Software, die Evaluation und Zusammenfassung unserer Erkenntnisse. Außerdem wird mögliche weiterführende Forschung besprochen.

Acknowledgements

There are, as always, far too many people I would need to thank, but I will still try.

First of all I want to thank my supervisor Gero Herkenrath for not only being an exeptional supervisor, but also introducing some much needed calmness in situations of time pressure. Thank you Gero. I really do appreciate it.

I also thank the whole Media Computing Group and Prof. Dr. Jan Borchers. I really enjoyed working with all of you and appreciate the input and the open doors you provided. Thank you.

I would like to thank the University of Edinburgh for the amazing year I had studying there and especially Dr. Mark Wright for accepting second examiner without ever meeting me in person. I sincerly do thank you Mark.

I also want to thank the city of Aachen for their consent to what proved to be a valuable study. Especially the support of the nice staff of the city hall in Aachen made the study a unforgetable experience, even in times when nothing worked as expected or at all.

A very special thanks goes out to my friend and collegue Florian Herlings who did not only help me debug tedious c code, but also encouraged me to go on until this very document is done.

An unusual thank you is dedicated to Locas in Love, a band from cologne that I listened to while wrapping up my thesis. Amazing work guys. Thank you.

Finally, I would like to thank my parents Dagmar and Norbert and my brother Malte: Mum, Dad, I could have never done this without your neverending support and encouragement. I think you were more worried than me. But guess what, I just finished my thesis. Malte, you are just the most awesome brother one can wish for. Thank you.

Conventions

Throughout this thesis we use the following conventions.

Text conventions

Definitions of technical terms or short excursus are set off in coloured boxes.

EXCURSUS:
Excursus are detailed discussions of a particular point in a book, usually in an appendix, or digressions in a written text.

Definition:
Excursus

Source code and implementation symbols are written in typewriter-style text.

```
myClass
```

The whole thesis is written in British English.

Download links are set off in coloured boxes.

File: myFile[a]

[a]http://hci.rwth-aachen.de/public/folder/file_number.file

Chapter 1

Introduction

Comprehension of human behaviour has long since been a goal in the area of psychology, sociology and human-computer-interaction since their initial foundation were layed out. New business branches emerged over the past decades only dealing with human behaviour and what one can learn from this behaviour, eventually using these insights in marketing, design, architecture and also in the computing industry.

One noticable application is seen in museums trying to understand how visitors perceive their exhibitions, how they move around in their venues, what makes them stay longer or keep more of the content in mind. Ultimately this leads to the question: How can visitors be engaged?

One possible answer to this question is the use of multimedia or audio guides in order to supply the visitor with content that is tailored to certain exhibits or rooms. These audio guides used to be a device, that had a numeric keypad for entering the numbers next to the exhibits. Afterwards the audio guide would play an audio recording for the exhibit in question.

Nowadays these guides have become more sophisticated. There is a lot of effort put into context-aware guides that either react to exhibits or rooms automatically or by user interaction e.g. QR Codes next to exhibits. The automatic approach usually facilitates indoor location tracking or sensor techniques placed close to exhibits. These guides show contextual information without the interference of the visitor.

The city hall in Aachen has such a guide, named Aixplorer, in place. It has been developed by the Media Computing Group of RWTH Aachen University and provides the visitor with contextual information on a room-level.

The question that comes to mind is: Do those multimedia guides really influence visitor behaviour and engage visitors more than just the museum would? This is the central question that I want to answer in my thesis. The basis for my work is the Aixplorer and the city hall in Aachen.

In the following paragraphs, I will give a short summary of the chapters that you can find in this thesis.

2—"Related work" gives an insight on work that has been done analysing movement patterns as well as what these results can mean for guides. Furthermore, I will look into methods of capturing user behaviour.

3—"User Study" explains the details of the user study that has been constructed to explore visitor behaviour with and without Aixplorer in the city hall in Aachen. This chapter also shows the camera setup that I used to observe the study participants.

4—"Software Design" outlines the specifications that the software evaluating the results needs to implement, including the annotation of video material collected during the study and the analysis of the resulting annotation data set.

5—"Software Implementation" shows the features of the implemented software and shall serve as a manual for the use of it.

6—"Evaluation" gives an overview of patterns and behaviour that has been observed and possible interpretation of these.

7—"Summary and future work" shall finally sum up the work that has been described in my thesis and give an outlook on possible future research based on the data I collected.

Chapter 2

Related work

In the following I will look into relevant work that has been done in the areas of movement pattern analysis, adaptive guides and tracking methods. Furthermore, I will outline the features of the multimedia guide "Aixplorer" that is the subject of our study. The movement pattern analysis is interesting for us, since there have already been studies to show how people move around in galleries without any multimedia content. Also there has been a study regarding guides that adapt to user behaviour and demography, which could be a valuable use case for the data I collect during our study.

The tracking technologies and methods I looked into are merely camera-based solutions as this proved to provide the best compromise of cost-efficiency and invasiveness during the study. I also looked into tracking technologies using infrared or heat cameras, as well as movement sensor and kinect-based solutions. However, these solutions were either too expensive to implement or not suitable for the partially large rooms in the city hall in Aachen.

2.1 Movement patterns

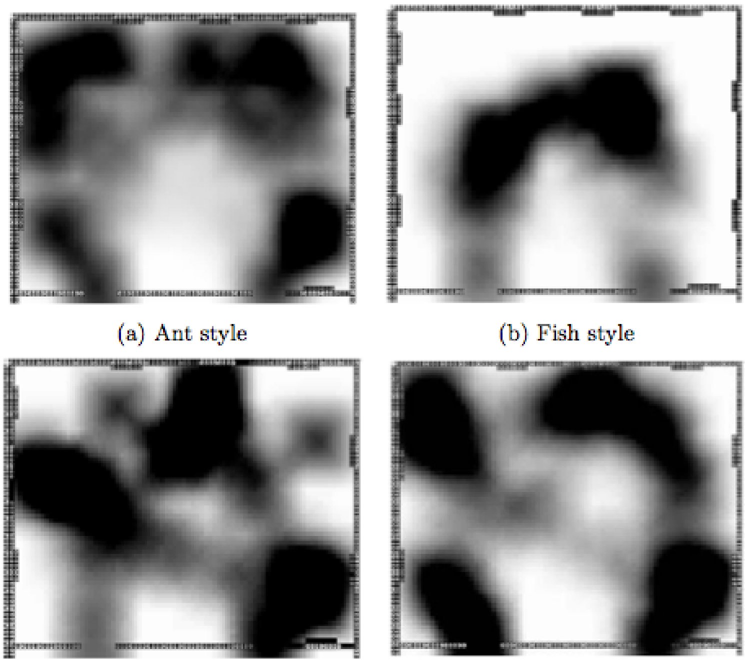
The tracking of movement patterns has been done several times in the past to analyse visitor behaviour in museums.

Most of the publications available refer to a book published by Véron et al. [1983]. The author identifies four distinct behaviour styles of visitors. Those four have been entitled using animal names: ant, fish, butterfly and grasshopper. The ant visitor will follow a distinct path and tries to look at each and every exhibit there is, while the butterfly visitor is rather guided by the physical location of the exhibits than a path, but also stops regularly to get detailed information. The fish visitor can be found in the centre of the exhibition rooms most of the times and avoids observing details of the exhibits. Finally the grasshopper tends to have a specific field of interest and only observes certain exhibits. However the grasshopper spends a lot of time observing the ones that interest her, while ignoring the others. One visitor can present one or more of those patterns during one visit.

2.1.1 Synthesizing visiting styles

Sookhanaphibarn and Thawonmas [2009] and Chittaro and Ieronutti [2004] propose ways to visualise the four visiting styles. Chittaro et. al. visualise the visitor behaviour based on the time the visitor spends at a certain point in an art museum. Despite the fact that the authors were researching movement patterns in a virtual environments, they claim that the findings can also be a applied to the physical space.

Sookhanaphibarn and Thawonmas [2009] refined that approach since the Chitarro methodology was created for an art museum and all exhibits are placed on the walls, meaning that there were no exhibits that were placed elsewhere in the room. The authors propose four approaches to analyse and mathematically express each of the four visiting styles based on Veron's definition, instead of using one methodology for all of them. They build a mathematical model including four formulas to visualise the movement data of visitors. These visualisations, to be found in figure 2.1, are a reasonable suggestion to visualise the data they collected and their style also inspired my heatmap visualisations.



(a) Ant style (b) Fish style

(c) Grasshopper style (d) Butterfly style

Figure 2.1: Visualization of visitor styles excerpted from A Visual Tool for Tracing Users Behaviour in Virtual Environments by Chittaro et.al. in 2004 - Excerpt by Sookhanaphibarn et. al.

2.1.2 Restrictions

The initial research by Veron was done in a museum that was not using any kind of multimedia guide, but further research in this area supports their thesis for multimedia guide usage as well and proposes a facilitation of the identified visiting style, proposing methods for adaptive systems that correspond to the visiting style [Bianchi and Zancanaro, 1999, Zancanaro et al., 2007]. This means, that the user is presented with different content based on her preferences.

All of those studies including Veron's however either do track visitors wandering around the museum without any guides or with a multimedia guide. There is no direct comparison between the two types of visitors in place. I will close this gap, by analysing both visitors with multimedia guide and those without multimedia guide.

2.2 Multimedia guides

2.2.1 Aixplorer

The Aixplorer is in use at the city hall in Aachen

The Aixplorer [Herkenrath and Borchers] is a multimedia tour guide developed by the Media Computing Group in cooperation with the city of Aachen in 2009. Is is currently in use at the city hall of Aachen.
Apart from supplying the user with multimedia content, it allows for location based tracking in the city hall. Essentially the user is confronted with multimedia content that has been tailored for the room she is currently in. Technically this has been realised using Wi-Fi spots within the city hall. The use of the Wi-Fi beacon, in particular the SSID/-mac address and the calculated signal strength allows for tracking the user over two floor levels and determine the room she is currently in.
Also a Ubisense [1] installation is in place in the biggest room, the coronation hall. In this particular room, tracking of the visitor is more accurate. It tracks the head orientation of the visitor to supply her with audio material that is both adjusted in volume the closer she comes to the source as well as it makes localisation of the source possible through individual volume adjustment of both earpieces.

2.2.2 Adaptive audio guides

The study conducted and documented in "A User-Centered Approach to User Modeling" by Petrelli et al. [1999], analyses the possibility of adaptive content in audio guides. The "HyperAudio" guide that is already in place serves as a model example of how they would go about adapting their approach. During their study the authors

[1] Ubisense is a proprietary location tracking system - http://www.ubisense.net/en/

asked visitors to answer questions on five ”topic areas”. Namely:

- *Personal data profile*, which included basic demographical data.
- *Museum habits*, asking for the frequency and visiting preferences (in a group, alone)
- *Context of the current visit*, which covers the motivation
- *Course of the current visit*, identifying the use of guides (humans, as well as exhibition labels)
- *Styles of visit*, focussing on the behaviour

These ”topic areas” resulted in a set of four user model variables, that were created and used to configure the audio guide before it is handed to the visitor. These variables are summed up as ”Family, school or adult” (group behaviour), ”First-time visit”, ”Foreseen visit duration”, ”Interaction preferences”.

The findings above are a valid proposal for the usage of our extracted data. Not only can it be used to explore vistor behaviour during the visit, but also the insights collected can be used for adding adaptivity to the Aixplorer; ultimately using a more user-centred approach to guiding different visitors through a museum.

Since an adaptive system is not my first priority, I do not want to collect all the data neccessary from the visitor. However future work could make similar suggestions based on the data I collect.

2.3 Tracking technologies

As already mentioned above, I looked into various methods of tracking people through technologies, like infrared cameras or movement sensors. Apart from the tracking methods based on cameras, there was a second possible solution founded in the room-based tracking of the guide itself. In the following I will outline the different possible tracking solutions we can use tracking visitors.

2.3.1 Aixplorer WIFI tracking

The Aixplorer facilitates the wireless infrastructure of the city hall in order to determine the room the user is currently in [Herkenrath and Borchers]. The advantage of this tracking technology is that the location data is closely connected to the user's behaviour. It allows logging of events such as multimedia guide use that can be directly related to the user behaviour. i.e. a user does not move for 5 minutes, since he is listening to an audio comment targeting exhibits in the very room she is currently in.
Another upside of this kind of tracking is that it is minimally invasive and the observation would not affect the behaviour. However the inbuilt wireless tracking is inaccurate enough to only allow for room based tracking right now, which would not give me the granularity I am looking for.

Definition: *Room-based tracking*

ROOM-BASED TRACKING:
At the writing of this thesis the current Aixplorer Version is running on an iPhone Original, a version for the new iPhone 4S was in preparation. However the room-based tracking is really accurate with a fairly small error margin.

This would mean that I am only able to determine the room visitors are in and not the exact location within the room which would result in very inaccurate data, that does not give us the necessary resolution we are looking for i.e. Heatmaps for movement within rooms are not possible.

Also one more critical disadvantage of this tracking method is that I can only track people with an Aixplorer. Those participants who do not use an Aixplorer can not be tracked with this particular method.

2.3.2 Camera-based methodology and algorithms

Video-camera-based methods and algorithms have a clear advantage over other techniques (heat cameras, movement sensors, infrared). Cameras are not only relatively cheap to come by, but are also available in fairly small scales e.g. action cameras or in mobile phones.

During my research I came across a number of security systems and research in people surveillance which were able to track people very accurately. However the techniques were neither open-sourced nor accessible for academic use. e.g. Calderara et al. [2008]. That is why I looked into general people recognition algorithms.

There are various ways of identifying humans in a video. I have evaluated three algorithms and approaches in preparation of the experiment conducted. In the following these approaches will be presented and commented on regarding their feasibility for our use-case.

Most of the approaches presented below facilitate automatic body detection for tracking visitors around the museum. This can be done by installing cameras in each room and then process the collected data through e.g. body or blob detection.

OpenCV HOG

The Histograms of Oriented Gradients for Human Detection (HOG) is an approach proposed by Dalal and Triggs [2005]. Their basic idea was based on the finding that the appearance of local objects and shapes can be expressed through the "distribution of local intensity gradients or edge directions". The exact position of the gradients or edges is not important for this approach. In practice the image is first of all colour- and gamma-normalised, the implementation then divides the image into cells (containing x pixels) and calculates a 1-D histogram of gradient directions or edge orientations. These histogram entries are then combined to form the representation of the image. Ac-

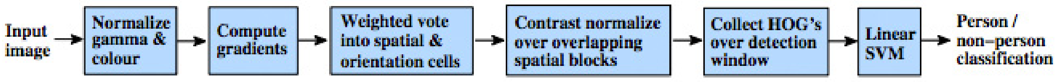


Figure 2.2: HOG Pipeline

cording to the implementation chain (see 2.2) the cells are contrast-normalised, which is done by accumulating multiple cells to a block, then taking the measure of local histogram energy and using the results to normalise the cells. Those created descriptor blocks are called "Histogram of Oriented Gradient(HOG) descriptors". Finally the detection window is being tiled with a "dense (in fact, overlapping) grid of HOG descriptors". The resulting feature vector is used in a Linear SVM to classify humans / non-humans.

KSP

K-shortest path tracking is an algorithm that has been developed in the CVLab at the Ecole Polytechnique Federale de Lausanne [Berclaz et al., 2011, Fleuret et al., 2008]. It enables more precise multi-object tracking and could be facilitate for my experiment as well. The KSP is part of a pipeline consisting of several independent tools / programs, some of which are not open-sourced by the authors. The pipeline works as follows:
The videos recorded need to be split into their frames and are passed through a background subtraction algorithm. Those results are forwarded to the POM, which is a people detector. The latter results are batched together and passed to the KSP algorithm. The KSP tracks them together and passes the results to the identifier, which contains identity information for people distinction. e.g. number at the back of player jerseys or colour of clothes. After a close evaluation and further information retrieved from the developing department at the Ecole Polytechnique Federale de Lausanne, this method is not feasible for my experiment. The usage of this algorithm would result in a high quality of data, but the effort and preparation time exceeds my capacities. For future experiments the pipeline presented above should definitely be looked into.

POM stands for Probabilistic Occupancy Map

OpenTLD

The OpenTLD (former Predator), as presented in "P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints" [Kalal et al., 2010] is another approach to automated tracking. It allows for manual selection of an object to be tracked and uses P-N learning (Positive-Negative learning) to continuously track the selected object. P-N learning is an approach that exploits both the labelled and unlabelled data to track the object. Furthermore the algorithm is iteratively learning from positive and negative constraints based on the structure of the unlabelled data. This results in the classifier getting better, the longer the object is tracked. i.e. it can not only track the face but is, after learning from the video data, also able to track the back of the head.

This approach is tremendously interesting for me, since people in a museum cannot only be tracked by their faces, since they turn around every-so-often. However during some initial testing this algorithm did not perform as well as expected especially with difficult light scenarios i.e. bright lights or dim lights.

2.4 Summary

I have shown, that there has already been efforts to visualize visitor paths in museums [Sookhanaphibarn and Thawonmas, 2009, Chittaro and Ieronutti, 2004] and also potential methods of using user profiles and visiting behaviour for adaptive guides [Bianchi and Zancanaro, 1999, Zancanaro et al., 2007].

These approaches have, however, not used cameras for accurate user tracking and therefore we had to discuss other approaches [Dalal and Triggs, 2005, Berclaz et al., 2011, Fleuret et al., 2008, Kalal et al., 2010] to analyse video material. Ultimately none of the algorithms or processing pipelines discussed above did give me the needed accuracy without further optimisation for my use-case. One reason is the lighting situation in the museum. Some rooms are very bright, some are fairly dark. Also I need to follow people through multiple rooms and therefore tracking one visitor would have meant processing 12 videos. With over 20 visitors that would have resulted in multiple days of post processing. Due to time constraints, I therefore opted for a manual annotation approach that will be discussed in 4—"Software Design".

All in all, especially the visualisation of different visiting styles and methods how to use the collected data, shall give an insight on how we can facilitate the results of our study, that will be presented in 3—"User Study"

Chapter 3

User Study

The historic city hall in Aachen is one of the most prestigious sights in Aachen and is therefore a popular visiting place for tourists. The museum has been equipped with location-aware multimedia guides by the Media Computing Group of RWTH Aachen University [1]. These guides are room-sensitive and supply the visitor with content tailored for the very room they are in (see 2.2.1).

In the following, the experiment conducted at the City Hall Aachen as well as the goals, requirements, design, implementation and preparatory process involved will be outlined.

3.1 Goals

This section covers the goals we persue in the study. These are ″Understanding visitor behaviour″, ″Recommendations for exhibitions and guide design″ and ″Observation of visitor paths″, whereas the second one will not be covered in this thesis since it requires a deeper analysis of the user behaviour and shall therefore be left for future work.

[1] http://hci.rwth-aachen.de/

3.1.1 Observation of visitor paths

By constructing this study I want to get a high resolution visitor path for the city hall in Aachen for the first time ever. Until now, only a room-based tracking approach could be implemented with the use of the Aixplorer. Also only visitors with Aixplorer could be tracked, which I will now change by implementing a camera-based approach.

3.1.2 Understanding visitor behaviour

The results of the study shall help to understand how visitors behave in the museum and especially whether they behave differently when they use the Aixplorer. Also I want to understand what influences visitor behaviour e.g. do visitors in a group behave differently from visitors that explore the museum on their own. These can be done by looking at the data we extract from the videos as well as consulting the videos itself for a qualitative analysis.

3.1.3 Recommendations for exhibitions and guide design

A third goal that inevitably emerges when exploring the visitor behaviour is: What do the results mean for exhibitions and guide design? Although I will not make any recommendations in this thesis, the data can help museums, in particular the city hall, to improve the design of their exhibitions e.g. exhibit positioning, visitor flow. Also my results can be useful for the improvement of multimedia guides (see 2). e.g. content recommendations based on demography or user behaviour.

3.2 Requirements

In order to run the experiment I had to fulfil a distinct set of preliminaries. I needed consent from the city of Aachen, since the museum is also part of the current city hall and used for council meetings. Also, the choice of appropriate cameras for the historic location was of importance. In the following I will outline the above and give insight on our decision process.

3.2.1 Privacy

Although most people are monitored by Cameras everyday e.g in supermarkets, stations, airports and many other public places, the implications for privacy regulations need to be looked at very closely. As mentioned above, the city hall in Aachen is not only a museum, but also plays an active role in political life in Aachen. The mayor and the city council as well as other city employees work in the same building and the mayor regularly welcomes guests there. After consulting with the privacy officer and the major and supplying them with a description of the study (A.4,A.5,A.6), they could not find any concerns as long as the following conditions are met during the experiment:

- Every visitor needs to agree to be filmed even if she does not want to participate in the experiment.
- In case a visitor disagrees with her being filmed, the cameras need to be shutdown immediately.
- Every participant needs to sign the consent form A.1.
- The camera footage cannot be publicised. Only the extracted movement data can be used in publications.

3.2.2 Demography & Background

In order to get some demographic background on our study participants I need to create a basic questionnaire that covers age, gender and profession. Also I will access the technical background and their experience with both the city hall and the multimedia guide (see A.2). That way I can evaluate whether demography and experience influences the visitor's behaviour.

3.2.3 Camera selection

For the selection of appropriate cameras for the locations I had to think about a few factors. First of all the image quality needed to be good enough to visually identify people for the manual tracking process. Also the building itself restricted the choice of potential cameras. As the museum is a historical building, I was not allowed to permanently modify the building or use any measures that could potentially damage it. There is no possibility of connecting the cameras to a power supply or link them in a globally accessible network. Furthermore I needed the possibility of triggering as many cameras as possible at the same time with some kind of remote trigger mechanism. Another softer criteria was the size of the camera. Since people should not be intimidated by the cameras or feel watched, I was looking for small cameras.

I finally decided to go with GoPro Hero 3 White Edition [2]. These action cameras fulfil most of our requirements very well, since they are not only small and light, but they can be run using batteries. Also the integrated wireless functionality can be used to connect up to 50 cameras to one remote, which makes simultaneous recording possible. The last thing that is crucial during camera setup, is a preview image. This can be retrieved by connecting the goPro to an iPhone (via wireless) using the official goPro App [3]. The only problem is connected to the battery time. The batteries only last for 2-3 hours depending on the camera resolution and standby time. The activated wireless adds to the power consumption. For that particular reason the Media Computing Group bought a second set of batteries that replaced the used ones after half of the study day.

The cameras come with cases, that allowed me to temporarily mount those cases in the museum and switch the cameras (daily setup and battery switch) and yet maintain the same position throughout the experiment. (For camera positions see 3.3.2)

[2]Specifications can be found in A.8 and A.9

[3]AppStore: https://itunes.apple.com/app/gopro-app/id561350520

3.2.4 Camera positions

After selecting the cameras I needed to think about the positions of the cameras in the museum. I need to capture as much of the room as possible and in order to do so, I need to place the cameras as high as possible. Placing them under the ceiling pointing towards the floor was due to architectural restrictions, not possible. That is why I had to think about the next best thing, which is mounting the cameras at the highest point possible, pointing down to the floor. Most of the walls in the city hall contain ledges, that make such a setup possible.

The requirements adumbrated above are implemented in 3.3.2—"Camera positions".

3.3 Design & Implementation

The City Hall in Aachen has 7 rooms accessible for visitors. The lobby, the council hall, the white room, the red room, the master craftsmen's kitchen, the master craftsmen's court, the coronation hall and a staircase. For this experiment we excluded the council hall and the staircase. The council hall is often not available for visitors, since it is used by city hall officials for city meetings. The staircase is not feasible for camera monitoring since it would require another 5-6 cameras in order to get a proper result. Therefore the lobby, the 4 adjacent rooms and the coronation hall (first floor) are monitored, resulting in a total of 6 rooms. In the following I will outline the terminology used, the decisions made for the camera positions as well as the preparation plan and the study flow. Also I will explain why I decided to use room markers for the accurate marking of the room size.

3.3.1 Terminology

In order to understand the setup of the study, the results and the evaluation, I need to explain the terminology of the setup first.

Whenever I refer to room or camera numbers (1-9) these correlate to the rooms shown in 3.1—"Camera-Room Assignment".

Camera Identifier	Room
1	Lobby
2	White Room
3	Master Craftsmen's Court
4	Master Craftsmen's Kitchen
5	Red Room
6	Coronation Hall (bottom-right camera)
7	Coronation Hall (bottom-left camera)
8	Coronation Hall (top-left camera)
9	Coronation Hall (top-right camera)

Table 3.1: Camera-Room Assignment

The exact positions, including identifier, of the cameras are explained in 3.3.2—"Camera positions" and are are shown in Figure 3.2 and 3.6.

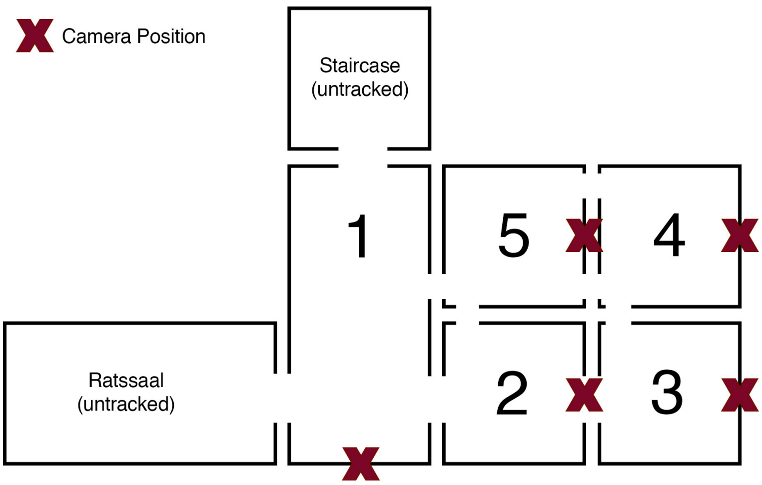


Figure 3.1: Ground Plan: Camera positions for the ground floor. Cameras are marked with a red cross

Figure 3.2: Photos: Camera positions for the ground floor. Cameras are marked with a red circle.

Figure 3.3: This is an example for camera position tests I did in the white room. It gave me insight on the angle and position as well as the necessary height of the cameras.

3.3.2 Camera positions

I first explored the possible camera positions by taking test pictures (see 3.3). As the test picture shows, I was initially experimenting with positioning the cameras somewhere in the room corners. However, I discarded the idea of doing so. The reason being the lobby and the coronation hall (first floor). Due to architectural restrictions, I was bound to place the cameras at one of the room sides (see 3.1 and 3.5). A homogeneous camera setup simplifies the analysis of the video material collected, that is why I decided to setup cameras in the other rooms accordingly. Figure 3.2

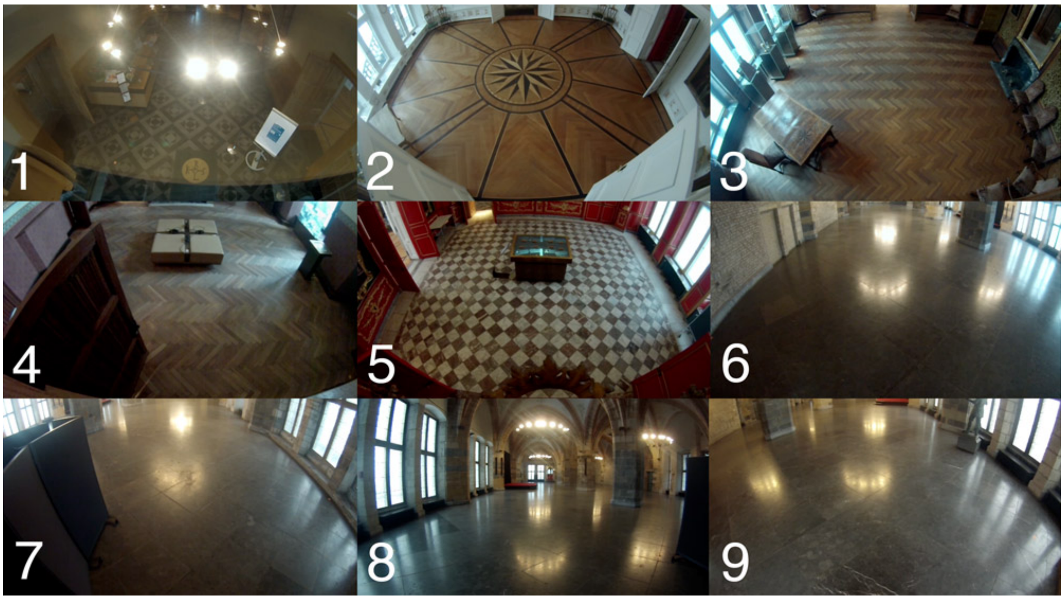


Figure 3.4: Camera Viewports

and 3.6 show exactly where and how high the cameras have been mounted. These two figures in combination with 3.1 and 3.5 give a complete overview of the setup.

As you can see in Figure 3.2 and 3.6 the cameras have been mounted as high as possible to get a good angle of the room. With Camera 6,7,8 and 9 in the coronation hall this was not possible on account of architectural restrictions. After reviewing test images of those cameras, the height of the cameras still proved to be satisfactory. The cameras 1 through 5 have been mounted adequately high and give a good angle for visitor observation. An overview of all camera viewports can be found in 3.4

To maintain the same position throughout the experiment the cases were glued to the according positions using blue tack and Tesa power strips for traceless removal after the experiment.

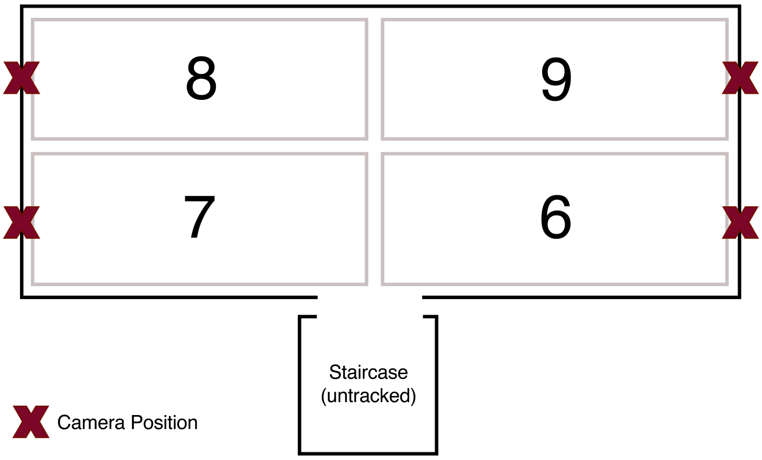


Figure 3.5: Ground Plan: Camera positions for the first floor. Cameras are marked with a red cross

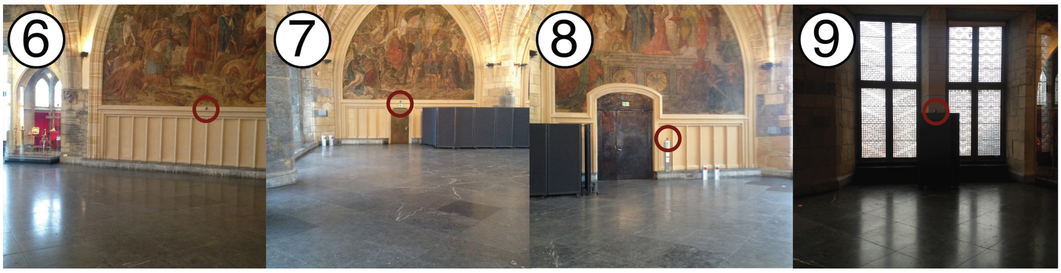


Figure 3.6: Photos: Camera positions for the ground floor. Cameras are marked with a red circle.

3.3.3 Room markers

Because I need to be able to accurately calculate the ground truth for the tracking path of a visitor (see 4), there are 4 ground plan markers for every room (16 for the large coronation hall/first floor). These markers ensure that the room is the same size on every video during annotation. The reason why I use these markers is that the rooms do mostly have round corners. In some cases these markers reduce the room size minimally. Since this is consistent with every single visitor, it does not influence the results of the study.

Figure 3.7: An Example for the markers we put in every corner for every camera tracked area

3.3.4 Study flow

On entrance into the museum visitors are presented with a poster that notifies them of the ongoing study and expresses what we intend to do with the data we collect from the study (Figure 3.8). (see 3.2.1—"Privacy").

I explain the study to them and ask for their participation and permission to use their data. The consent form shown in A.1 outlines the exact workflow of the study and the util-

Figure 3.8: Poster informing about the study and its implication

isation of the collected data. The visitor is talked through the consent form, signs it and gets a copy signed by myself for reference [4]. Afterwards she is asked to fill in a one-page questionnaire A.2 for later demographic evaluation of the study.

Since the privacy regulations were outlined in the consent form A.1 and explained by myself before the experiment,

[4]Consent form also contains contact details, in case visitors change their mind after the study. - However this did not happen.

no participant was uncomfortable with being filmed, even if they did not want to take part in the study. The consent form contains a clause, that the visitor consents with being filmed but disagrees with the usage of her data. This has been a preliminary of the city Aachen authorities (see 3.2.1—"Privacy").

In case the visitor decides to take an Aixplorer with her, it is prepared and explained by museum staff. The tracking starts after all the formalities are done and the visitor is on her way (with or without Aixplorer). As a compensation for taking part, the entry fee and Aixplorer renting costs of every participant are paid by the RWTH Aachen.

The study flow as presented also meets the other requirements by city officials that were delineated in 3.2.1—"Privacy".

3.3.5 Daily preparation

For each day of the user study, there is a set of preparatory steps necessary. These steps are outlined in the preparation form A.7

Day before preparation

First of all, the batteries (numbered 1-24) and the remotes (R1, R2) need to be charged. Secondly, the SD cards need to be emptied, if filled from the day before and the video material is archived. All the Batteries are kept outside the cameras until the study begins to avoid loss of energy due to stand-by.

On day preparation

Once at the museum the batteries are inserted and the connection to the remotes is setup based on the camera position[5]. Afterwards the cameras are inserted into the cases, that have been mounted in the museum. Once this has been completed for every camera, the remotes are checked again. If it works, the poster informing about the study is setup and the forms are all prepared.

3.4 Results

Overall I had 41 participants of which I can use 21 for my evaluation. Due to camera failure during the experiment the video material captured for other participants was not usable (see 3.4.1).

In the following I will summarise the study's resulting data. The interpretation of this data can be found in 6—"Evaluation"

3.4.1 Difficulties

During the experiment I experienced some difficulties using the cameras (goPro Hero 3 white edition). When I connected the cameras to the according remotes, cameras would randomly crash during the experiment. Sometimes after 2 or 3 minutes and sometimes after 1 hour. There was no obvious pattern recognisable.

This led to the significant number of study participants I cannot use for the evaluation. The reason for this is that video files were destroyed during crashes and therefore there was no usable material for certain participants.

3.4.2 Demography

The demography of this study has been determined through questionnaire A.2. 62% of the participants were female, 38% were male. On average the participants were

[5] one remote for the first floor, one remote for the ground floor

36 years old. Visitors using the Aixplorer were on average 42 years old, visitors without the guide on average 31 years old.

Group	Participants
Group 1	PN1, PN2, PN3, PN4
Group 2	PG2, PG3, PN5, PN6, PN7, PN8, PN9
Group 3	PG5, PG6
Group 4	PG7, PG8
Group 5	PN10, PN11

Table 3.2: Visitor groups

3.4.3 Experience

Also the experience of participants has been accessed through questionnaire A.2. Only one of the participants had visited the museum before, 13 out of 21 participants have used a multimedia or audio guide in a museum before. The average visitor did have a smartphone and classified her smartphone experience as average.

3.4.4 Groups

Most of the participants were part of a group when they visited the museum, only 4 were on their own. 3.2—"Visitor groups" shows the group constellations.

3.4.5 Video recordings

The video recordings obtained during the study serve as the basis for our evaluation. Since the videos need to be run through the annotation process and analysis first, these results will be discussed in 6—"Evaluation".

3.5 Summary

I have shown how I planned and executed the study, what I intend to learn from it and showed the first results. These results will be discussed in 6—"Evaluation".

Chapter 4

Software Design

The chapter Software Design will cover the design decisions that have been made for the process of annotation and analysis. The primary goal of the Software to be designed is extracting the visitor paths from the video data and get it into an appropriate structure. This structure is then used to visualize and analyse the data. I will in particular answer the following questions:

- How do the tasks for the two processes look like?
- What is the right platform for either processes?
- How does the data transfer between the two processes look like.
- Is the annotation and analytical process part of the same tool, or do we split these two processes and create two tools?

4.1 Requirements

The first thing I will look at is the required functionality for both, annotation and analysis.

4.1.1 Annotation

The annotation process will be done manually, since the automated methods discussed in 2—"Related work" did not work for our scenario / would have resulted in post-processing annotation. That is why I opted for a manual process in the first place. Manual annotation must be done anyhow, since also automatic algorithms make errors that need human intervention.

For our annotation process this means that we need an interface that can perform the following tasks:

1. Create rooms and assign the recorded videos to the individual rooms.

2. Trigger a new tracking process for every new visitor.

3. Follow the participant in the museum, by switching rooms while annotating.

4. Tracking a visitor using a pointing device (must be convenient since we have 21 participants to annotate)

5. Export of the tracking data (In case we opt for a two tool solution)

6. Fast forward functionality: In case you need to skip certain parts of the video e.g. when visitors enter the staircase.

7. Rewind functionality: In case you make errors while annotating

8. Marking the room dimensions on the video, in order to calculate the ground truth.

9. Input of room data, including actual room size measured between the room markers

4.1.2 Analysis

The analysis also has a set of requirements as I will show below. The software needs to perform the following tasks on the resulting data of the annotation process.

1. Calculate the ground truth
2. Show ground plan of the museum
3. Load single participants
4. Show visitor paths for single participants
5. Playback functionality: Path over time.
6. Show statistics for single participants
7. Show heatmaps for single participants and groups e.g. with / without guide, in order to visualise hot spots.

4.2 Choice of platform

I choose the iOS and Mac OS X as the platforms to look at. First of all they are available for all major application types: smartphone, tablet, desktop applications and the Media Computing Group primarily works with Apple products. This makes iOS or Mac OS X the appropriate choice, since the tools will most likely be used for future research as well.

4.2.1 Annotation

I have shown the tasks the tool has to perform. We will now look at the individual tasks in order to decide on an appropriate platform. Most of the tasks could be performed fairly well on most of the platforms that have been outlined in 4.2—"Choice of platform". Since we are annotating videos, the device I annotate the video on, should have a large enough screen to properly mark the points in the video the visitor is at. That is why I discarded a smartphone application immediately. That leaves two more options: Either a desktop application or a tablet application will be implemented.

For all but one of the tasks that the tool needs to implement, both platform would, in my point of view, perform equally well. However the task: "Tracking a visitor using a pointing device" is the most important one and also the one that will perform better on a tablet device using a stylus.

First of all the annotation process will take a lot of time and tracking a visitor using a stylus on an iPad can be done practically anywhere. If we would opt for a desktop application, tracking would most likely require a mouse, since a touchpad is not accurate enough or a graphics tablet. This solution would not be as portable as an iPad app. However a mouse will be more accurate, since you can clearly tell what the cursor points at. A stylus for an iPad has a more fuzzy focus. In a short feasibility test, I tried using a mouse for the annotation and realised that it is hard on my hands after a while. This is the major downside of a desktop-based application. You can argue that you can also use a stylus for annotation in a desktop-based application. However I do not have a touchscreen, but would have to use either a drawing tablet or a stylus on the touchpad which results in a similar fuzziness as with the use of a stylus on the iPad.

Despite the mouse being the more accurate tracking device, I still decided to implement the annotation application as an iPad app. The overall handling (stylus annotation) makes the annotation process more enjoyable, which is in this case a valid argument, since the annotation will take roughly 25 hours and the impact on my hands is reduced immensely. Also after testing the stylus method I realised that the focus of the stylus is not as fuzzy as expected and therefore qualifies for an annotation process.

4.2.2 Analysis

For the analytical part there is again a larger display necessary to appropriately display the ground plan of the museum. That is why a smartphone is definitely too small. One could visualize one room at a time, but I decided that I want to see the whole ground plan at once, since it makes the analysis more convenient.

Again I could break it down to either a tablet or desktop application. In this particular case, I cannot see any reason not to implement it for either platforms. However I would like to have a large interface in order to have all rooms visualised at once. The iPad [1] does give me a good resolution (2048x1536), but the display is with 9.7 inches still significantly smaller than a desktop or laptop display. This is why I opted for a desktop application for the analysis. This decision makes it necessary to discuss a data interface in 4.3, since we now have two tools on two different platforms.

4.3 Data interface

I described the decision process that led to the making of two tools for two different platforms. This leads us to a different problem that needs solving: How does the data format and transfer look like?

4.3.1 Data requirements

The analytics application needs a certain set of data from the annotation application. First of all the data regarding the rooms need to be transferred. This data includes the size that has been measured between the markers (see 3.3.3—"Room markers") and the room size, that has been marked on the video.

The second data set that the analytics application needs, is

[1] iPad specifications - http://www.apple.com/ipad/specs/

the position of a participant at a certain point in time. This includes the room, the exact timestamp and the pixel position in the video.

4.3.2 Data format

The two different data sets described above will be distributed into two separate logfiles. These logfiles can be exported for every participant. The reason for this is that the marker positions can vary in the videos due to camera and battery exchanges. Therefore the export contains two files:

The roomNumbers 6 - 9 refer to the coronation hall, since there are four cameras in place

- rooms.csv - which contains the actual room size in the museum and the exact coordinates of the rooms within the recorded video. The format of the rooms.csv file is as follows roomNumber, top_real, right_real, bottom_real, left_real, top_video, right_video, bottom_video, left_video.

- tracking.csv - which contains the visitors tracking results, including the room she is in and the exact coordinates marked on the video. The format of the tracking.csv file is as follows: camera-/roomnumber,timestamp,x_position,y_position.

4.3.3 Data transfer

The two csv files need to be transferred from the iPad to the desktop application. I decided to attach the csv files to an email, that will be composed within the iPad app. That way I have the possibility of commenting on a specific participant, if there is any unusual behaviour. The subject will be the participant identifier.

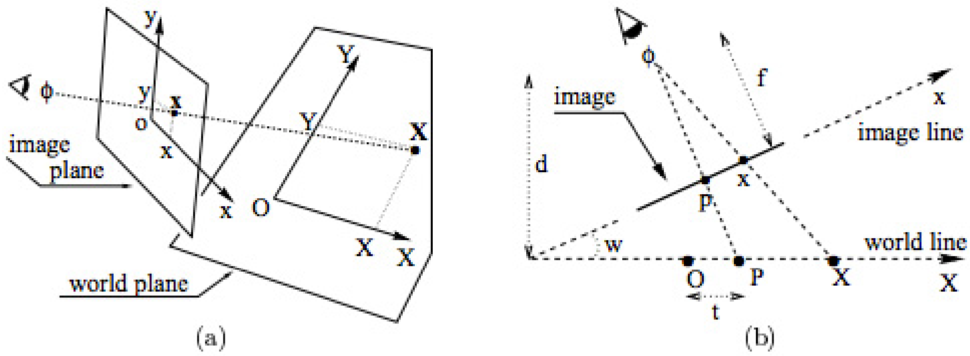


Figure 4.1: Excerpt from "A plane measuring device" by Criminisi et al. [1999]. The point x(x,y) in the image plane is transformed into the point X(X,Y) in the world plane. The image plane is the position in the video, the world plane the actual position in the museum.

4.4 Ground Truth Calculation

A very important part of the tasks described for the analytics tool is "Calculate the ground truth". This is why this part of the tool will be covered in more detail. The path of the visitor that has been marked during annotation needs to be transformed into the actual path in the museum.

4.4.1 Perspective Transform Estimation

The algorithm picked to perform the task has been developed by Criminisi et al. [1999] at the University of Oxford, published in "A plane measuring device".

Originially the Algorithm has been developed to measure uncertainty in images. Applications named by the authors include "furniture placements and interior design purpose".

In my approach the algorithm is used for the ground truth calculation which is in this particular case a perspective transformation as shown in 4.1. Based on Criminisi et al. [1999], Christopher R. Wren wrote a companion on the implementation of the algorithm published on the MIT Website [Wren, 1998] that I will use as a reference for my implementation.

In his companion the author implements a Matlab solution, that transforms a free-form rectangle (image plane) defined by the user into a square angled rectangle (world plane). Furthermore the user is able to select two endpoints of a line, that are assigned to the image plane and then transformed into the world plane. 4.4—"Ground truth calculation" (a).

4.4.2 Algorithm

The findings of Criminisi et al. [1999] result in a formula that is used by Wren [1998] and that works as follows:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 & -X_nx_n & -X_ny_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -Y_nx_n & -Y_ny_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ X_n \\ Y_n \end{bmatrix}$$

Figure 4.2: Excerpt from Wren [1998]. x1 - xn and y1 - yn referes to the edges of the image plane. X1 - Xn and Y1 - Yn refers to the edges of the world plane. The vector λ (a - h) holds the results for transforming any given point P (image plane) to P′ (world plane)

The edges of the image plane rectangle and the edges of the world plane rectangle are filled into the equation isolated by Criminisi et al. [1999]: 4.2—"Ground truth calculation equation". The form of the equation is Ax = B and can be solved as shown in 4.3—"Solution for the Ground truth calculation equation". For more details as well as solution process, please confere Criminisi et al. [1999].

$$
\begin{aligned}
A\lambda &= B \\
A^{\top}A\lambda &= A^{\top}B \\
\lambda &= (A^{\top}A)^{-1}A^{\top}B
\end{aligned}
$$

Figure 4.3: Shows the solution for the equation in 4.2—"Ground truth calculation equation". (see Wren [1998]). λ holds the results needed for the necessary transformation.

4.4.3 Proof of concept

I took the Matlab implementation by Wren [1998] and transformed it into a proof of concept for the required functionality and as a testing tool for the implementation to be done for the analytics tool. Contrary to C. R. Wren's implementation, I used a fixed source rectangle and one fixed source point in my image plane to be transformed into the world plane. This coheres to the actual calculation that needs to be done in my tool. The source code can be found in Software Design & Implementation: B.1—"Perspective Transform Estimation Matlab Code". An exemplary output is shown in 4.4—"Ground truth calculation" (b).

4.4.4 Limitations

The algorithm described above performs accurately when there is no distortion inflicted by the camera lens. This is however usually the case, which is one of the reason why Wren [1998] refers to his implementation as "Perspective Transform Estimation". There is always small distortions and in our case the camera uses a wide-angle lens and therefore there is distortion, especially at the rim of the image. A much higher accuracy would be achieved calibrating the cameras for every time we set them up again. I still decided to use the algorithm explained in 4.4.1—"Perspective Transform Estimation".

The reason for this is that time did not allow for the effort that I would have to put into camera calibration and ground truth calculation. The accuracy I get from using "Perspective Transform Estimation" is high enough to get reasonable results. This becomes evident when looking at

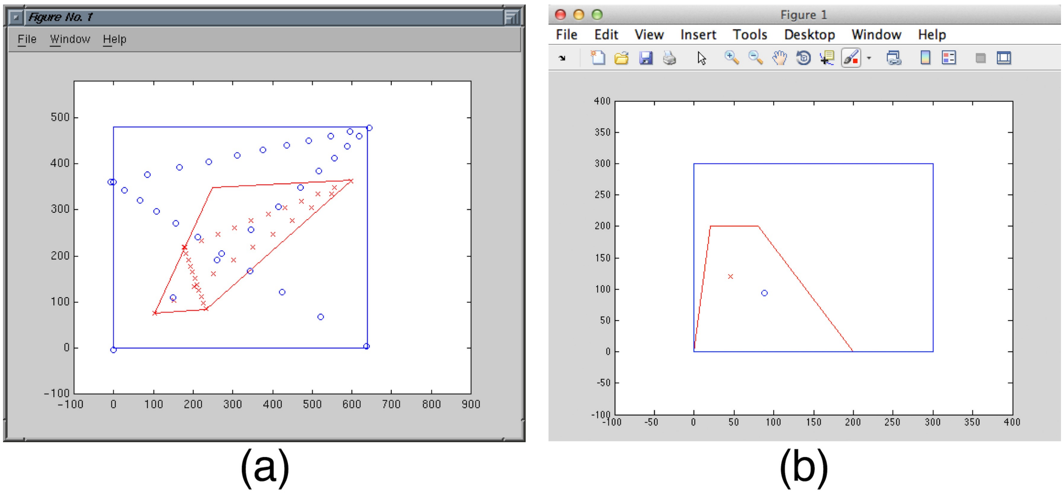


Figure 4.4: (a) Excerpt from Wren [1998], showing a Matlab implementation for mapping a line from the image plane onto the world plane. (b) shows a Matlab implementation for just mapping a point from the image onto the world plane. This is my implementation based on Wren [1998]. Code can be found in B.1—"Perspective Transform Estimation Matlab Code".

the results we achieved and especially the evaluation chapter of this thesis (see 6—"Evaluation")

4.5 Workflow

The following section shall give an overview of the exact workflow that leads to the visualisation of the movement data and brings the requirements defined (see 4.1 Requirements) into a process. Also the steps necessary to use the video data will be outlined.

4.5.1 Video processing

The videos that the goPro Cameras produce, are split into 1.9GB files. This is probably due to the fact, that the micro SD Card are fat formatted. In order to annotate the videos on the iPad the files are merged into one video file per room and compressed. The compression makes sure that the iPad memory is not exhausted, since 1 hour of video material is about 3GB in it's original resolution and quality. e.g. the first recording is 1:30h long, which makes 4.5GB per room, resulting in 40.5GB of data. The iPad, I use for annotation only has 32GB at its disposal and therefore the compression is necessary. Also it significantly lowers the load of the iPad during annotation.

4.5.2 Annotation application

The compressed video files are loaded onto the iPad using iTunes file sharing [2]. Afterwards we create the tracked areas [3] and enter the room measures of every single tracked room. This step is necessary to project the movement tracked onto the room layout. Secondly the recorded videos that were uploaded onto the iPad are assigned to the rooms. The tracking process will allow for marking the room edges on the video spanning a rectangle on top of the video. For every new video and room this process has to be repeated. This saves me the trouble of positioning the cameras on the exact same spot every day, which is hardly possible since also the cases installed might move during the exchange of cameras.

The visitor is tracked using a pen to ensure accuracy. As long as the pen touches the iPad the video is played and the coordinates of the touch are tracked. Once the pen is lifted, the video stops. Since there might be multiple participants in one recording, there is a fast forward functionality to jump to the starting point of their visit. Also if you make an error, the rewind functionality allows to jump back 5 seconds in time. The tool outputs two logs files (see 4.3.2 Data format) that are send via email in order to load them into the analytics tool.

[2] iOS: About File Sharing - http://support.apple.com/kb/ht4094

[3] As a reminder: For the coronation hall we need 4 tracked areas, which are handled like rooms.

4.5.3 Analytics application

The collected data supplied by the annotation tool needs to be visualised appropriately in order to analyse it. The logfiles are first of all imported and labelled by their participants. These logfiles can then be analysed in three different ways.

The logfiles can be visualised by showing the visitor's path to observe the visitor's way through the museum. This visualisation can be slightly altered to get a second one. By displaying the path over time you should be able to reconstruct the visit. Therefore the tool contains a playback functionality for the path. The third visualisation is heat map based. Heatmaps can be used for a single participant, but also for participant groups i.e. All, with Aixplorer, without Aixplorer.

Furthermore there is a statistics window displaying informations about the time a visitor spent in a certain room / area and how long her whole stay took. Also the exact room path can be found. e.g. PN1 1,5,4,3,2,1,6,9,6,1 (see C.1).

4.6 Summary

In this chapter I explained the decision process for implementing two tools (annotation and analytics) and the tasks each of those tools has to fulfil. I covered the exact workflow for the annotation and analytics process as well as the data interface to transfer the data from the annotation to the analytics tool. Also a proof of concept for the ground truth calculation algorithm has been presented, which is an important part of the analytics tool. In the following chapter I will show the implementation of the requirements outlined in this chapter.

Chapter 5

Software Implementation

This chapter will deal with the implementation of the Requirements I have defined in 4—"Software Design". I will cover the iPad annotation tool: FollowThem and the analytics tool: TrackingAnalyser. The focus of this chapter will be on the user interfaces of both tools and the details on the implementation.

The documentation on how to install and use the software can be found in D—"CD Content (Software & Documentation)".

5.1 Annotation application: FollowThem

FollowThem is the project name for the annotation tool, that enables me to annotate the recorded and processed videos. This section shall cover all relevant aspects of this app and will explain the interface design decisions that have been made.

5.1.1 Environmental setup

The first screen that pops up, once the app has been started is shown in 5.1—“Room details” (a). The left part of the screen shows all rooms that have been created so far and also serves the purpose of creation and deletion of rooms. Apart from the room label, every table cell includes the current video file, that has been assigned to the room.

A room can be edited by selecting it on the left side, which shows the details on the right side. This pattern is typical for iPad Master-Detail apps which use an `UISplitViewController`.

Definition: *UISplitViewController*

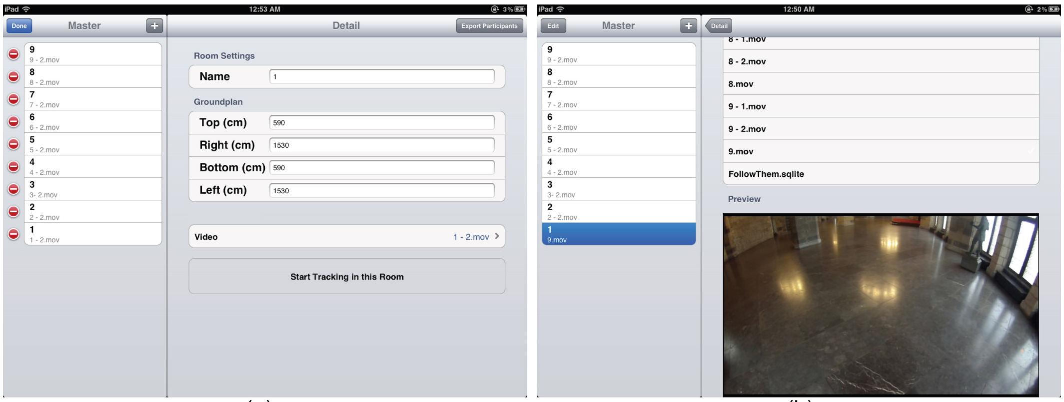
UISPLITVIEWCONTROLLER:
Master-Detail application on the iPad use a UISplitViewController. This Controller holds the Master Controller on the left-hand side, usually showing an item overview or a navigation. On selection of an item on the left-hand side, the right-hand side of the screen (Detail Controller) is updated according. The Detail Controller, as the name suggests, show the content of the selected item / provides details for the selected item.

The right-hand side of 5.1—“Room details” (a) shows the three attributes (attribute-groups) that can be edited. The Name of the room serves as an identifier. The real measures in cm for the top, right, bottom and left side of a room will be used to calculate the ground truth (see 4.4—“Ground Truth Calculation”) and the video file assigned to the room can be chosen by selecting the table cell holding the video file name.

Selecting the video file brings the user to a new screen, replacing the Detail area of the UISplitViewController. 5.1—“Room details” (b) shows all the video files that have been uploaded to the Documents folder of the FollowThem iPad app. The upload is done via the iTunes File Sharing Interface as shown in 5.3—“iTunes File Sharing”.

The FollowThem.sqlite file contains the core data structure of the app and is also stored in the Documents folder of the app

5.1 (b) also shows a preview window for the currently selected video file. The preview helps choosing the right



(a) (b)

Figure 5.1: iPad App: Room details

video file in case the file names are ambiguous.

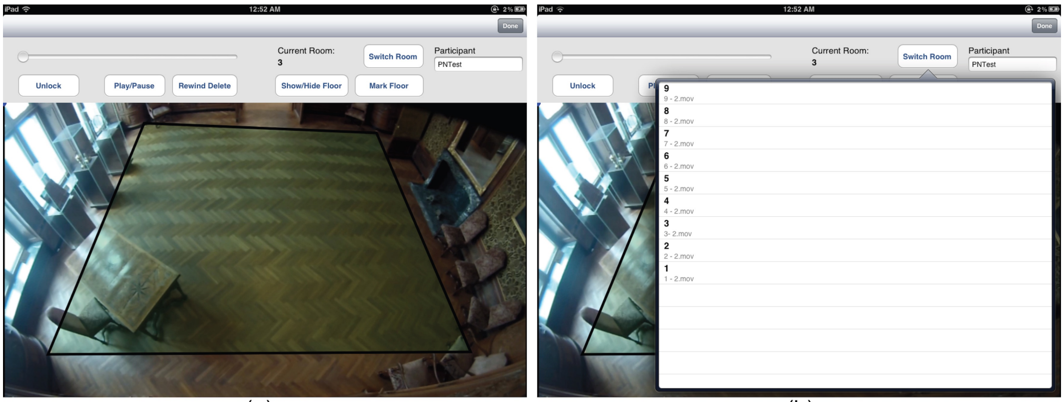
Once the video file has been selected the user can go back to 5.1 (a) and start the annotation process in the first room by selecting "Start Tracking in this room". This opens up a new Screen that I will turn to in the next section 5.1.2—"Annotation"

5.1.2 Annotation

The annotation screen 5.2—"Annotation" (a) is split into two parts. The top part of the screen holds the controls that are needed for a smooth annotation process, the bottom part shows the video of the selected room and the area that has been marked as the room (see 3.3.3—"Room markers").

Annotation controls

The control panel at the bottom can functionality-wise be split into two parts. The left parts shows the video controls including the slider to speed up the video, the pause / play button and the rewind delete button. The rewind delete button rewinds the video by 5 seconds and deletes all annotations that have been made during these 5 seconds. This



(a) (b)

Figure 5.2: iPad App: Annotation

functionality comes in handy, when you loose a participant while tracking her. The last button that needs explanation is the unlock button. This button has been added in order to lock / unlock all other controls in the panel. It shall prevent accidental presses in the controls panel, while annotating.

The controls displayed on the right-hand side of the panel allow meta data manipulation including the ”Mark floor”, ”Show/Hide Floor”,”Current Room”, ”Switch Room” and ”Participant”. ”Mark Floor” triggers the room marking process. The user can now select the four edges of the current room (see 3.3.3—“Room markers”). In case you do not want to see the marked room during annotation, you can hide it by pressing the ”Show/Hide Floor” button.

The Participant Input takes the label of the current participant that the user annotates and once the visitor leaves the ”Current Room”, the ”Switch Room” button can be pressed to shows the available rooms. The pop-up that shows the available rooms is shown in 5.2 (b). By selecting the next room the visitor walked into, the video for the current room is replaced by the video for the next room.

The room switch, does not reset the current timestamp. e.g. A visitor walking from Room 1 into Room 2 at 1:34, makes the user switch from Room 1 to Room 2. The video of Room 2 starts playing again at 1:34 thus ensuring a seamless annotation process.

Annotation window

The bottom part of the screen 5.2 (a) not only displays the video for the current room, but can show the tracked area of the room (green rectangle with black border). Also the whole video area is responsive to touch input for the tracking process. I will hereby explain how this exactly works.

Whenever the visit of a participant starts, the user starts pointing a stylus or one of her fingers to the feet of the participant. On touch the video starts to play and the user can track the participant by moving the stylus along the walking path of the visitor. Once the stylus is lifted, the video stops. This allows pausing the tracking process and also makes the room switches easier.

A stylus is highly recommended for the tracking process, since the accuracy is better than using your finger.

Once a visitor walks into the next room, this room can be selected in the control panel (see 5.1.2) and the video of the according room is shown in the annotation window. The tracking can now be continued by touching the video area with the stylus again.

While the stylus touches the video area, a background thread logs the current position of the touch, the current room and the timestamp into a database. This interval can be configured in the source code of the FollowThem app. For my annotation this interval was set to 0.2, which means that the background process logs the location and time information 5 times per second. For higher accuracy this interval can be set to a lower value.

It is not recommended to use the fast forward functionality during annotation, since this alters the logging interval by factor video speed e.g. interval is set to 0.2 and the video speed is set to 10 would only log the visitors position every 2 seconds.

When the annotation is done, the "Done" button closes 5.2 (a) and the user returns to 5.1—"Room details" (a).

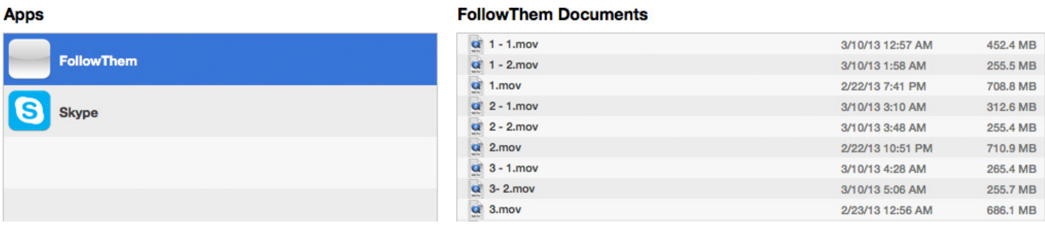


Figure 5.3: iTunes File Sharing

5.1.3 Data interface

The application does not only need an interface for exporting the participant data, but also one for the import of video files. The next two sections are dedicated to describe the two parts of the data interface.

Import

As already mentioned in 5.1.1—"Environmental setup", I enabled iTunes File Sharing for the application FollowThem. iTunes File Sharing is a convenient way to upload data to the Documents folder of an iOS app. The data can be uploaded when the iOS device (in our case: the iPad) is connected to iTunes. 5.3 shows the responsible part of the iTunes interface, that can be found in the app section of our connected device.

As for the FollowThem app, the user can select video files from her local harddrive and iTunes syncs them into the Documents folder of the FollowThem application. The video files that have been synced can now be accessed by the FollowThem and used for video annotation.

Export

The second part of the data interface is the export functionality for the annotation and room data collected for each participant. As described in 4.3.2 there are two logfiles (rooms.csv and tracking.csv) that need to be exported

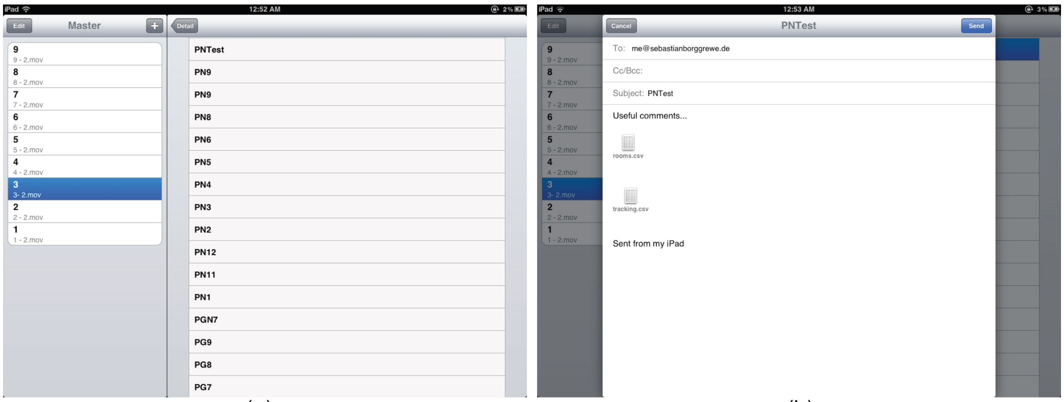


Figure 5.4: Participant Export

for every participant. These logfiles are generated from the data that has been stored in the local database of FollowThem. 5.4(a) shows the all participants that have been annotated.

Selecting a participant exports the data using a simple email functionality. (see 5.4(b)). The two logfiles are attached to that email, the subject is set to the participants identifier and the standard recipient is filled in. The loose format of this export allows the user to comment on specific participants for later use in evaluation.

5.2 Analytics application: TrackingAnalyser

The TrackingAnalyser is the analytics tool that shall help me understanding the user behaviour and compile a set of statistics and visualisations for me. In this section I will guide through the user interface and the implemented functionality of this tool and show examples for its usage.

The tool consists of two major components. There is first of all the control and statistics panel, which shows information about the participants and is responsible for triggering the visualisation. Secondly there is a ground plan of the museum that serves as a view for the visualisation of paths

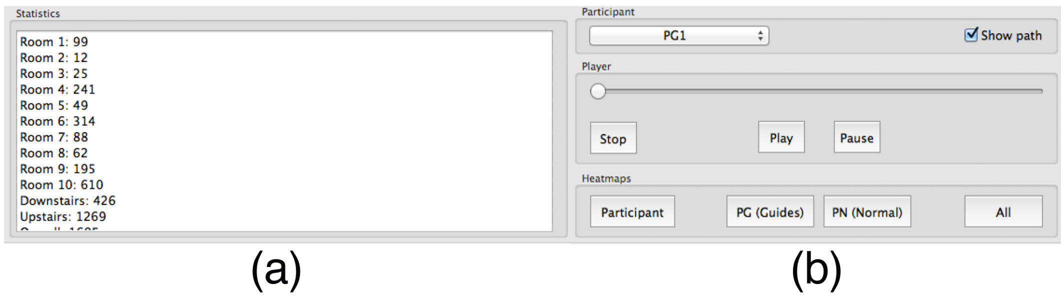


Figure 5.5: Control & Statistics Panel

and heatmaps. The latter two have each been dedicated a separate section due to their complexity.

5.2.1 Control and statistics panel

The Control and statistics panel 5.5 contains four sub panels: Statistics, Participant, Player and Heatmaps. All four of those will be covered in this section.

Statistics

The very left sub panel 5.5(a) contains information about the current participant. This includes the time spent in rooms and certain areas of the museum (also staircase) and the order of rooms visited. The data extracted from this panel can be seen in C—"Evaluation"

Participant

The participant panel holds a list with every participant of the study. The control shown in 5.5(b) shows the currently selected participants. All[1] information and actions in other sub panels refer to this participant. The second control in this panel "Show path" indicates whether the path visualisation shown in 5.6 is switched on.

[1] All but the heatmap panel. The heatmap panel also holds options that apply to a group of participants

Player

The player panel does also show the user path, but it does so over a fixed period of time. The path for the selected participant is visualised in the same manner as with ”Show path” (see 5.6). However the player is able actually show the visit in a configured number of steps (at the time of writing 100). This means that the overall time spent in the museum is split into 100 chunks of data that are presented to the user chunk by chunk.

Heatmaps

The heatmap panel is as shown in 5.2.1 different from the other three. Apart from the ”Participant” Button, the heatmap panel can trigger heatmap generation for more than only the selected participant. There is one button for all participants using the Aixplorer, one for all participants without Aixplorer and one for visualising all visitors in a common heatmap. An example for those heatmaps is shown in figure 5.7.

5.2.2 Visualisation

The explanation of the control panel functionality already referred to two different visualisation types within the analytics tool. These visualisations are generally speaking an overlay for the ground plan of the museum. This ground plan is always shown as part of the analytics tool.

While the first visualisation, the participant’s path, can only be shown for one participant at a time, the second one, the heatmap, can visualise multiple participants at once. In the following I am going to talk about the two visualisations and how they work.

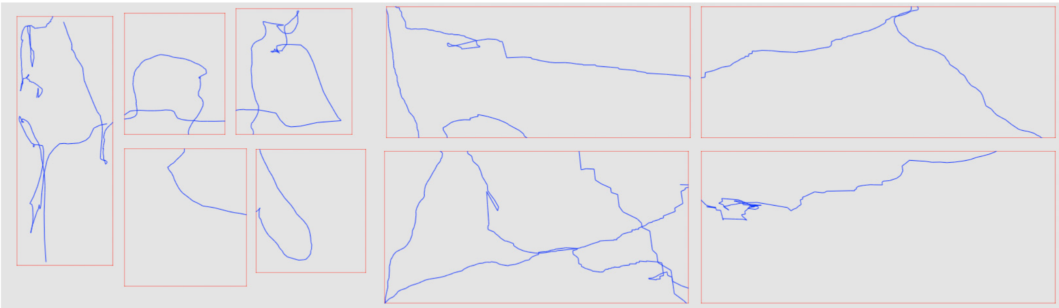


Figure 5.6: Path Example

Path visualisation

The visitor's path as shown in 5.6 does visualise the way she walked around in the museum. I have five locations per second and interpolate between them. This interpolation results in multiple Bézier curves, eventually forming a continuous line that represents the visitor's path.

The interpolation is done per room and there are two edge cases one has to think about, when implementing this visualisation. Namely:

- If there is a time difference that is significantly larger than 0.2 seconds the two adjacent points within a room cannot be connected, since the participant has most likely been to another room in between.
- If the distance between two adjacent points is larger than a certain plausible threshold[2], I don't connect the two points

This behaviour leads, in most cases, to a continuous path. There are however a few participants, where you can see small gaps in the path, which can either be explained by an annotation error or an accidental "rewind delete".

In case the player is used to display the path over time, the procedure is the nearly the same. The only difference is that the path is displayed bit by bit and not at once.

[2]A participant is very unlikely to run 10 meters in one second

Heatmap visualisation

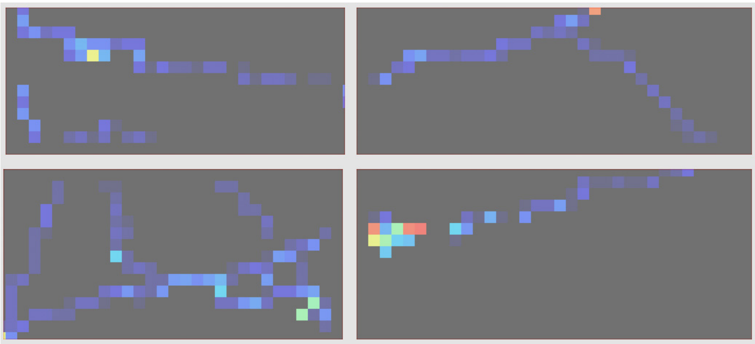


Figure 5.7: Heatmaps Example for the coronation hall

The tool's heatmap visualisation (see 5.7) can be done for multiple scenarios: One participant, only Aixplorer participants, only participants without Aixplorer and all participants. Nevertheless the heatmaps are produced in the same way.

First of all the room is subdivided into a grid that consists of 64x64cm large tiles. The tile size has been choosen according to visualisation within the tool: 4cm in the real world room are approximately expressed by 1px in the ground plan. According to this scale one tile measures 16x16px. 64x64cm is not only a convenient measure for calculating the grid, but also a reasonable size for a persons current location.

The visitor's locations are then filled into the grid. e.g. if a given location correlates to grid position 4-8 in room 1, 4-8 in room 1 is incremented by 1. All grids in all rooms are build like that.

When visualising the grids, the algorithm iterates through the grid and reads the value for every tile. For an appropriate visualisation, I defined six different colour ranges for specific tile values. The thresholds for different colours have been picked randomly at first and were then refined by learning from the visitor data and optimised for a reasonable visualisation of the visitor locations. The results of this process is shown in 5.1—"Heatmap colour of tile for given grid value".

Value of tile	*Colour*	*Description*
0	grey	No time spent in this tile
<5	shades of blue	under 1 second
<20	blue & shades of green	under 4 seconds
<50	shades of blue & green & shades of red	under 10 seconds
<100	shades green & shades of red	under 20 seconds
>100	shades of red	over 20 seconds

Table 5.1: Heatmap colour of tile for given grid value

If the heatmap visualises more than one participant, the value of the grid tiles is normalised by division of the number of participants. This method proved to give good visualisation results for our collected data.

An exemplary heatmap for the coronation hall is shown in figure 5.7.

5.3 Summary

In this chapter I discussed the software implementation, how it is used in order to annotate our videos (see 5.1—"Annotation application: FollowThem") and how I can analyse it (see 5.2—"Analytics application: TrackingAnalyser"). Furthermore, I have shown how the analysis tool visualises the data in heatmaps and paths and shown examples of those. In the next chapter 6—"Evaluation" I will make sense of the data that I gathered using the two tools.

Chapter 6

Evaluation

In this chapter I will discuss the results of the user study with the help of the analysis software that has been described in 5.2—"Analytics application: TrackingAnalyser". I will evaluate whether our hypothesis *H1* can be verified with the data collected:

H1: *Visitors with the Aixplorer behave differently to visitors without the Aixplorer*

The evaluation is based on observations that have been made using the data I obtained from the annotation and analytics tool e.g. time spent in rooms, but also contains qualitative observations based on the video material and made during the annotation process. e.g. group behaviour.

PNs and PGs:
PN and PG will be used in the following as an abbreviation for *P*articipant *N*ormal and *P*articipant with *G*uide

Definition:
PNs and PGs

6.1 Observations

6.1.1 Time spent in rooms

The table C.3 displays the time spent in the tracked rooms for each participant of the study. Also the overall time in the museum and in the pre-defined areas staircase, ground floor and first floor are shown. In the following I will explain the data.

Ground floor

The ground floor is the place, where visitors spent the most time while in the museum. However this time spent on the ground floor is unevenly distributed throughout the rooms. As 6.4—"Heatmap - all visitors - ground floor" clearly shows, visitors spent most of their time in Room 5 and 4. A comparison with 6.2—"Exhibits ground floor" explains that behaviour. The exhibit density in those two rooms is higher and also the exhibits in room 4 and 5 are, apart from one, interactive. Room 3 does also have three display cabinets, but since they are not interactive, people spent less time there.

Room 2 is a special case, among the ground floor rooms. PNs tend to just walk through it and the visiting time lasts less than 1 minute on average. PGs on the other hand spent nearly two minutes and over twice as much time in this room. In my opinion this has to do with the exhibit density of this room. There is no real exhibits except from one table and portraits. However these results are already a hint that the audio information supplied through the guide extends the visiting time in rooms.

Staircase

As already explained, the staircase has not been monitored in the study, but I can still determine how much time a visitor spent there. In the staircase the visitor can find interactive exhibits[1] and also a large window that offers a view at the cathedral of Aachen. In the time spent in the staircase one can see large discrepancies among the study participants. There are 5 visitors who spent about 2 minutes or less in the staircase, which is approximately the time you need to get upstairs if you do not pace. And there is 3 visitors that needed between 8 and 13 minutes. The average is just over 5 minutes. One observation that I made, while checking on the cameras in the coronation hall, is the fact, that this time does not only seem to be influenced by the interactive exhibits in the staircase, but also by the view you have at the cathedral of Aachen. Most visitors stopped to look out of the window to enjoy the view.

Coronation hall - first floor

The first observation to be made in the coronation hall (6-9) is connected to the left area, recorded by cameras 7 and 8. This is the part of the first floor where participants spent the least time in. This is due to the fact, that there is no apparent exhibits in this area. However you can still see a slight difference between PGs and PNs. Visitors using an audio guide in the coronation hall and discovered CORONA[2] still spend time in this area. For group 5 this time is even higher than for any other visitors and groups. The reason for this is that they discovered CORONA and also passed the guides to group members without an Aixplorer, leading to an overall longer stay of this group. As shown in table C.3 the PNs in this group can therefore not be regarded as PNs in the coronation hall, but must be evaluated as PGs.

Still, visitors spent most of the time in the right area of the room, recorded by cameras 6 and 9. This is, as already hinted above, connected to the exhibits in this area of the coronation hall (see 6.3) The overall time spent on the first floor is, probably due to less exhibits, less than the time spent on the ground floor.

[1] Touchscreen displaying information

[2] Corona is an interactive audio experience based on the Ubisense system. Futher explanations can be found in 7.2—"Glossary"

6.1.2 Common visitor paths

This section shall show the results obtained from the visitors route through the museum shown in C.1—"Order of visited rooms by visitor".

First room visits

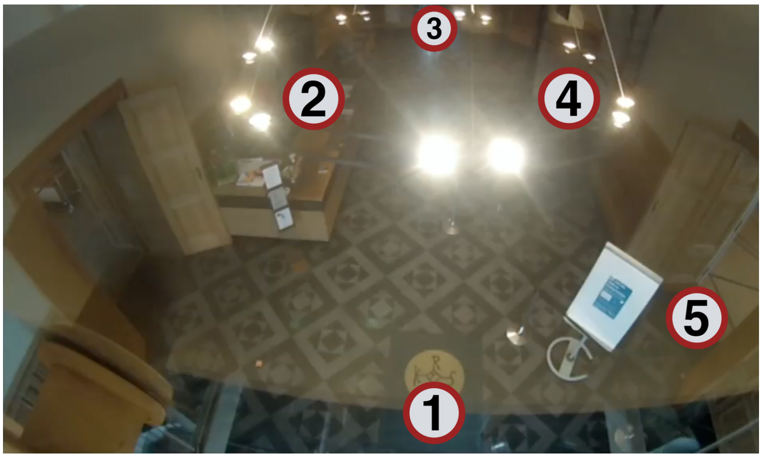


Figure 6.1: Lobby: (1)Entrance (2)Aixplorer Disposition (3)Staircase to first floor (4)Room 5 (5)Room 2

When visitors start their tour around the museum, there is one movement behaviour that is really striking. The way they choose their first room to visit (after the lobby) apparently depends on whether they use an Aixplorer or not. While 100% of all PNs went to the first floor first ((1) in 6.1), only 50% of the Aixplorer users walked for the first floor straight away ((4),(5) in 6.1).

In my point of view this has to do with the Aixplorer disposal ((2) in 6.1). Preparing the Aixplorer takes a few minutes and the participants have some time to look around in the lobby. This allows them to identify the other two doors (4) and (5) leading away from the lobby. When you just enter the museum and do not get an Aixplorer prepared, the first thing you see is the door leading to the staircase (2) and eventually to the first floor.

Also while disposing the Aixplorer visitors usually ask where to start. Although they tell visitors, that there is no fixed route to take, they show them that there is a ground floor. Also when visitors explicitly asked, the staff would say something like this:

It does not really matter where you start your tour. Most of the exhibits are on the ground floor. The first floor is mainly nice to see because of the atmosphere (and the copy of the crown jewels).

Also there was usually a gesture attached to the statement pointing out the rooms on the ground floor. This combination of statement and gesture could have an effect on the visitor as well in their choice of path.

Anticlockwise movement

One very interesting observation one can see in the path data I collected is that 14/21 participants move around anticlockwise through the rooms on the ground floor (rooms 1-5). Only 5 visitors moved through rooms 1-5 in a clockwise manner and with 2 participants, there is no real pattern noticeable. Also the anticlockwise movement usually looks as follows: 1,2,3,4,5... e.g. PG7 or ...1,5,2,3,4,5... e.g. PN4 depending on the movement before. The first pattern can be observed when visitors visit the ground floor rooms first, the second, when the come from the first floor. One attempt to explain this behaviour can be found in the door sizes. The doors separating room 5 and 2 and the ones separating room 5 and 4 are larger than the one separating room 5 and 4. Therefore you already get a good look into room 2 even before you enter it. With some participants there are small variations in the pattern described above, that will be covered in the next section Exploration pattern

Exploration pattern

Another pattern that can be observed with participants that are part of a group (see 3.4.4) is what I call the exploration pattern. There were five groups participating in the experiment and in four of these groups this pattern could be observed in at least one of the group members. With the largest group consisting of seven participants, there were even three group members showing exactly this pattern. However this last group is the exception and might be related to the fact, that the group was big enough to be split into sub-groups.

The exploration pattern is characterized by repeatedly walking from one room into another and returning to the original room again e.g. PN3 ...3,*4,5,4,5*,1..., PG6 ...3,*4,5,4,5,4,5*,1, PG3 ...5,*2,3,2,3*,4...

This behaviour can be either explained by curiosity and the need to show other group members what they discovered or the fact that different people take a different amount of time to explore one room. The second case would then be a demonstration of boredom and a hint to other group members to move on into the next room.

Either way this behaviour is highly interesting, also because none of the single visitors presented the exploration pattern. It seems exclusively to be a group related behaviour.

Shortest path movement

In 6.1.2—"First room visits", I have shown how users pick the first room they go to. Based on those findings I will now show that users tend to choose the shortest path to get out of the lobby.

First of all, all visitors without audio guides went upstairs straight away (see 6.1.2—"First room visits"). The ones using an audio guides went upstairs first in 50% of the cases, 40% went to room 2 first and only one visitor, PG9, picked room 5 first. When visitors came down from the coronation hall they went from the lobby to room 5 in 100% of the cases.

In my opinion this behaviour can have two possible origins. The first one being, that people generally try to pick the shortest path. The second possibility is that visitors actively avoid the till area. The till is the place where the staff are based. I find the first option more likely, but have no solid data to prove either one of these claims. Analysing this behaviour will be future work.

6.1.3 Single visitors vs. groups

As already mentioned in 3.4.4 81% of the visitors were part of a group. Being part of a group changes the behaviour of the individual. One pattern that can be observed has already been discussed in 6.1.2—"Exploration pattern". However there are more observations to be made.

First of all all single visitors took an audio guide to be guided through the museum. This behaviour can probably be explained by the fact, that they have no one to talk to while they are visiting and are therefore open to the audio input the guide supplies.

Secondly, none of the single visitors presented the exploration pattern explained in 6.1.2.

Furthermore the single visitors stayed on average over 2 minutes longer than groups (see C.2). This time difference is not very large. However for the median the results look very differently. While the median for group visits is 1164 seconds, the median visiting time for single visitors is with 1944 seconds - over 14 minutes longer. The reason for these differences lies in two visitor groups: Group 2 stayed by far longer than any other group and all single visitors stayed at least over 2 times longer than Single Visitor 2. These two edge cases explain the difference between median and mean. The results suggest that on average single visitors will stay longer than groups.

6.1.4 Interesting exhibits

6.2 and 6.3 show a ground plan of the museum with important exhibits. The stars show statues, display cabi-

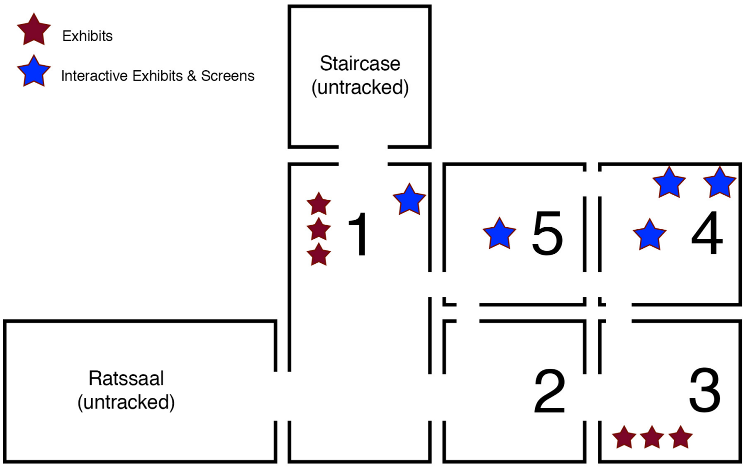


Figure 6.2: Shows the exhibits on the ground floor. This only includes statues, display cabinets, screens and interactive exhibits. It does not include pictures.

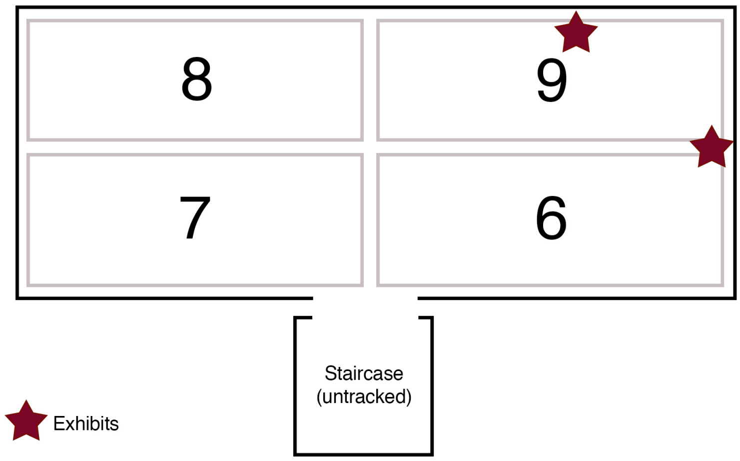


Figure 6.3: Shows the exhibits on the ground floor. This only includes statues, display cabinets, screens and interactive exhibits. It does not include pictures.

nets, screens and interactive exhibits. When I look at the heatmaps generated by our analytics tool, I can find a strong resemblance regarding those exhibits. As 6.4 and 6.5 show, the hotspot can be found around those exhibits shown in 6.2 and 6.3. In general interactive exhibits and screens were more engaging than others on the ground floor.

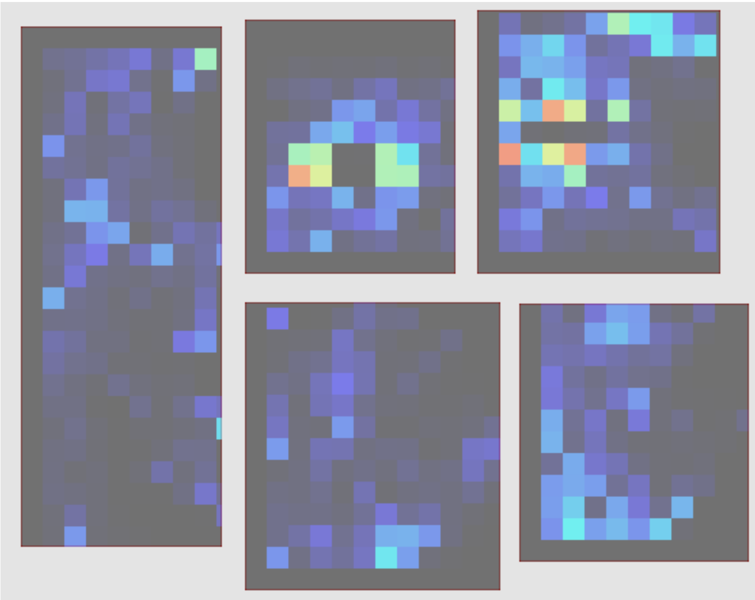


Figure 6.4: Heatmap showing time spent in all tracked areas of the first floor (all visitors)

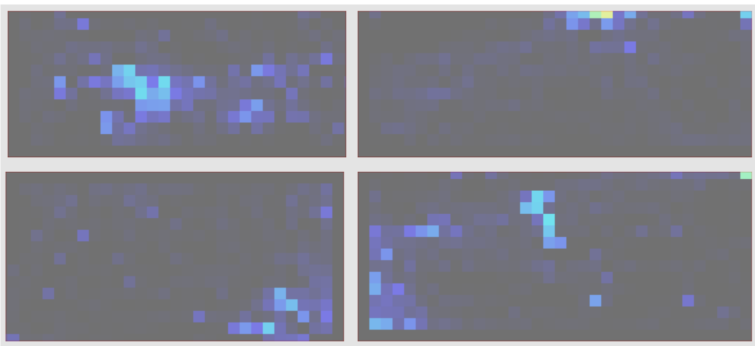


Figure 6.5: Heatmap showing time spent in all tracked areas of the coronation hall (all visitors)

The most looked at exhibits on the first floor were the statue and the copy of the crown jewels. While you can clearly see a hotspot around the statue (9 - top centre), there is not apparent hot spot around the crown jewels. The reason for this is that the cameras where not able to capture the area, where people were usually standing when looking at the jewels. The green spot at the very top right in 6 indicates that behaviour. Also I was able to clearly see that people spent time in this area while annotating the videos. This time should be measured in a future version of the analytics tool.

Figure 6.6: Room 3 - Camera View

Generally speaking there is a clear difference between multimedia exhibits and conventional ones. This is especially visible on the ground floor. The screens and interactive exhibits are more popular with visitors than the display cabinets or furniture displayed in the museum. Hotspots for wall mounted exhibits like pictures (excluding screens) cannot be found, except in room 3. As seen in 6.6, there is a mirror mounted at the centre of the right wall. This mirror does interest visitors. One group took pictures in front of them (group 1), but in general visitors just look at themselves in it.

The first floor only has two obvious exhibits and therefore those two are most popular and the higher popularity of multimedia exhibit does not apply here.

6.1.5 Resting behaviour

Room 4 is the downstairs room visitors spent most time in as already shown in 6.1.1—"Time spent in rooms". One of the reasons is the exhibit density (big screen, information screen, pictures, phones with audio material), the other is

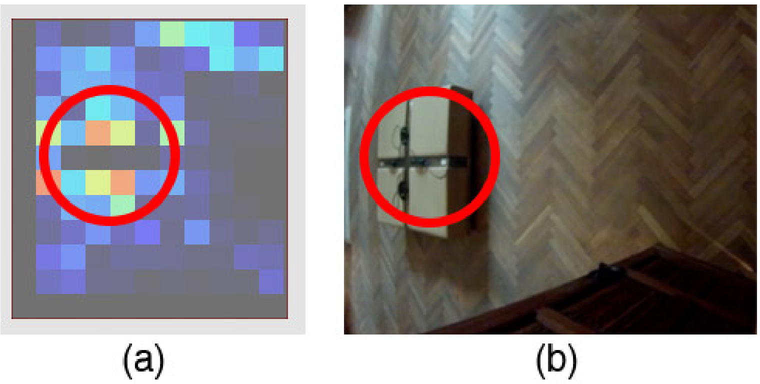


Figure 6.7: (a) shows the heatmap for all visitors for room 4, (b) shows the actual camera image (cropped and turned by 90 degrees to fit the room orientation)

definitely the four seating places located in the centre of the room. Most of the visitors sat down and looked around in the room or / and listened to the audio material.

However especially in groups people were using Room 4 as a waiting area for others to wrap up their visit or move to the next room.

Room 4 is the only room (except the lobby) with any seating, but also the room with the highest exhibit density. That is why I can't say for sure, that the time spent there is just related to the seating. However in my opinion, it adds to the time spent there, because of what I have observed in the video material i.e. visitors just sat around until other group members entered the room and then got up and left.

6.1.6 Passing around the Aixplorer

The participants marked as group 5 (see 3.2) is the only group in the study, that had PNs and PGs in it. This group therefore displayed a behaviour that could not be observed with any other group.

Once the two PGs realised the functionality of CORONA, they started handing the guides to the five PNs to try it for themselves. The group gathered at two points in the coronation hall and waited for everyone to try the Aixplorer.

Once the Aixplorer was handed over to the PNs, they started walking around in circles in the coronation hall. One participant was even pacing to get to the virtual audio sources.

This behaviour led to group 5 spending more time in the coronation hall, then any other group or single visitor. (see C.3).

6.1.7 Interactive table in room 5

Another observation I made during the annotation process is connected to the Interactive "Peace table". The table is a big screen, that reacts to IR sensors that are integrated into wooden blocks, that lie on the table. The wooden blocks are connected via a red line on the screen and display information about one aspect of the "Aachener Frieden"[3]

The blocks have 4 sides, that show different information on the screen for every side that touches the table. This possible interaction is however not perceived by the visitors. Only 1 out of 21 participants did realise that turning the block results in different information being shown on screen (PG5). He then showed it to his friend, he visited the museum with (PG6) and both started playing around with the table.

The two participants early spent 5 minutes in Room 5, more than any other participant in the study (see C.3—"Time spent in tracked rooms").

6.2 Summary

The Hypothesis, *H1*, defined at the start of this chapter will now be looked at again based on the findings described

[3]The "Aachener Frieden" is an historical event.

above.

Do visitors with Aixplorer show a different behaviour from visitors without Aixplorer?

As I have shown during the course of my evaluation, PGs did not only spend more time in Room 3, but their median visiting time is higher than visiting time of PNs. They also show a different behaviour in the coronation hall (6-9) if they discover CORONA. Mixed Groups (group 5 - PNs and PGs) even passed the guide around while in the coronation hall, in order to experience CORONA first hand. Also in 50% of the cases PGs choose to visit the rooms on the first floor first, none of the PNs exhibited this behaviour.

On a different note, all of the single participants in the study chose the Aixplorer over the unguided visit. Since I only had 4 single visitors, this should be confirmed by a study containing more single visitors.

All in all I have shown that *H1* can be verified, but I would still suggest to do the study again with adapted parameters. The reason is that "being in a group" also highly influences the visitors behaviour, as shown by e.g. 6.1.2—"Exploration pattern". I will suggest a different study setup in 7.2—"Future work"

Chapter 7

Summary and future work

The last chapter of my thesis shall offer a summary of my work and review the contributions made. I will also discuss suggestions for future work that should show the possibilities for further research and give the interested reader an insight on my expectations for the use of the data I collected.

7.1 Summary and contributions

The first contribution of this thesis lies in the software implemented to annotate and analyse video material in order to extract a user's path. This software can be used for future research and also be adapted to be used in other museums. The second contribution is the data set collected and extracted using the annotation software. I expect this data set to serve future research for extending the evaluations made in my thesis.

In 2—"Related work", I looked at studies that dealt with movement pattern analysis, the visualisation of those and tracking methods. Also I presented studies that encouraged the use of user behaviour and preference to supply adaptive guide systems.

After looking at the studies that have been done, in 3—"User Study" I closed the gap between studies doing movement pattern analysis for visitors with multimedia guides and those analysing visitor behaviour for visitors without multimedia guides. The study conducted resulted in the data of 21 participants that I could use for later evaluation.

In 4—"Software Design", I gave insight on the software necessary to extract the data needed for evaluation, including the annotation and analytics tool. Also I explained the algorithms needed for the ground truth calculation. The decision to split the tools into two parts for convenience reasons proved to be right for both annotation and analysis.

The implementation of those requirements were described in 5—"Software Implementation". I showed how the tools have been implemented and how to use them. Furthermore the technical details for the available visualisations of visitor data have been outlined.

6—"Evaluation" shows how the resulting data has been used. In this chapter, I reasoned with the data, being able to verify our hypothesis *H1*. I was able to extract visitor related patterns and observations from both the videos and the resulting movement data. The results have, due to time constraints, not been tested for statistical significance and most of them are qualitative observations. During the evaluation of the data I also realised that groups seem to influence the behaviour more than the Aixplorer does. Since I only had single visitors using the Aixplorer, I could not compare those with single visitors without any guide. However the findings serve as a first insight that is not only valuable for the museum, but also for the researchers working on the Aixplorer.

7.2 Future work

The repetition of the study with a different setup is the first thing, I would recommend for future work. The difference between PN and PG single visitors would be interesting to look at. In our study I could not make any assumptions about the difference between those two visitor types, because I only had single PG visitors. Also I would suggest executing the study in a different museum as well to find out which patterns and behaviour is due to architectural circumstances in the city hall.

The second thing that I would like to do is use the data I collected in order to create a framework for making the Aixplorer an adaptive guide learning from the demographical data of the visitor and the visiting style, resulting in the recommendation of user-centred content. However the kind of data necessary for implementing such a framework could not be collected during my initial study, since I did not collect enough information about the visitor. That is why I would recommend gathering more information about the visitor, when constructing a similar study again.

Also a good addition to the "adaptive" Aixplorer would be the possibility of live tracking a visitor through the museum and live-suggesting content based on the behaviour the visitor shows. This might seem utopian at the moment, but might be possible in the future.

The analytics software implemented in this thesis would benefit from further improvement as well. The statistical analysis should contain more metrics and visualisations i.e. group based tracking. Furthermore the automatic generation of statistics for all visitors or visitor groups is helpful for analysing behaviour and in combination with automatic significance tests on the data collected it will significantly improve the analytics tool.

Appendix A

User Study Material

Einverständniserklärung

Aufzeichnung von Bewegungsmustern im Rathaus Aachen

STUDIENLEITER Sebastian Borggrewe
Media Computing Group
RWTH Aachen University
Telefon: 017624873783
Email: sebastian.borggrewe@rwth-aachen.de

Ziel der Studie: Das Ziel der Studie ist es, zu verstehen wie sich Menschen mit Multimedia Guides in Museen bewegen. Die Teilnehmer werden gebeten, unter Anwendung des Multimedia Guides (Aixplorer), das Rathaus in Aachen zu besichtigen. Diese Besichtigung wird mit Kameras aufgezeichnet, um die Bewegungsmuster (gelaufene Pfade) erfassen zu können.

Ablauf: Sie werden von uns eventuell mit einem Multimedia Guide ausgestattet und besichtigen das Museum. Dabei werden Sie von unseren Kameras gefilmt, um Ihr Bewegungsverhalten aufzuzeichnen. Die Länge der Studie wird durch Ihren Museumsaufenthalt bestimmt und ist nicht beschränkt.

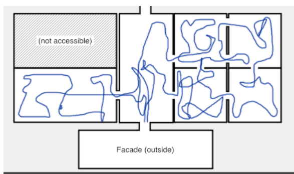
Risiken/Beschwerden: Es könnte sein, dass Sie die Teilnahme an der Studie ermüdet. Sie werden mehrere Gelegenheiten haben, sich zu erholen; zusätzliche Pausen sind ebenfalls möglich. Es sind keine weiteren Risiken im Zusammenhang mit der Studie bekannt.

Nutzen: Die Resultate der Studie sollen zum besseren Verständnis des Bewegungsverhaltens in Museen mit Multimedia Guides beitragen.

Alternativen zur Teilnahme: Die Teilnahme an der Studie ist freiwillig. Es steht Ihnen frei, Ihre Teilnahme zurückzuziehen oder abzubrechen.

Kosten und Entschädigung: Die Teilnahme an der Studie wird Ihnen keinerlei Kosten verursachen. Den Eintritt ins Museum und die eventuellen Kosten für die Bereitstellung des Multimedia Guides übernehmen wir für Sie.

Vertraulichkeit: Alle Informationen, die während der Studienphase gesammelt werden, werden streng vertraulich behandelt. Ihre Daten werden nur durch Nummern identifiziert. Keine Publikationen oder Berichte aus diesem Projekt werden personenbezogene Informationen über die Teilnehmer beinhalten. Wenn Sie sich bereit erklären, an dieser Studie teilzunehmen, unterschreiben Sie bitte unten. Die gemachten Bildaufnahmen von Ihnen werden ebenfalls nicht veröffentlicht, aus dem Rohmaterial werden nur die Pfade, die Sie im Museum verfolgen generiert, etwa in der folgenden Form (siehe rechts).



- ☐ **Ich habe die Hinweise auf diesem Formular gelesen und verstanden.**
- ☐ **Man hat mir die Hinweise auf dem Formular erklärt.**
- ☐ **Ich nehme an der Studie teil und stimme der Verwendung meiner Bewegungsdaten zu.**
- ☐ Ich darf während des Museumsaufenthaltes gefilmt werden, stimme aber der Benutzung meiner Bewegungsdaten nicht zu.

Name des Teilnehmers	Unterschrift des Teilnehmers	Datum
	Studienleiter	Datum

Figure A.1: Consent form for visitors that participate in the user study (german)

Angaben zu Ihrer Person

Vor der Teilnahme an der Studie, möchten wir Sie um ein paar Angaben zu Ihrer Person bitten. Diese Angaben helfen uns die Ergebnisse der Studie besser einzuordnen.

Bitte nennen Sie uns ihr Geschlecht?

[ ] männlich
[ ] weiblich

Wie alt sind Sie?

Ich bin ____ Jahre alt.

Was machen Sie beruflich?

[ ] Schüler/in
[ ] In Ausbildung
[ ] Student/in
[ ] Angestelle/r
[ ] Selbstständig
[ ] Rentner
[ ] Sonstiges

Haben Sie das Rathaus schon einmal besichtigt?

[ ] Ja
[ ] Nein

Haben Sie schon einmal einen Multimedia Guide in einem Museum benutzt?

[ ] Ja
[ ] Nein

Haben Sie ein Smartphone mit Touchscreen?

[ ] Ja
[ ] Nein

Wenn ja. Wie gut kennen Sie sich auf einer Skala von 1 - 5 mit dem Smartphone aus?

[ ] 1 - sehr gut
[ ] 2 - gut
[ ] 3 - mittelmäßig
[ ] 4 - schlecht
[ ] 5 - sehr schlecht

Figure A.2: User demography and smartphone usage questionaire (german)

Figure A.3: Poster informing incoming visitors about ongoing study (german)

Beschreibung der Benutzerstudie

Kurzbeschreibung
Wir würden gerne eine Studie mit **ca. 20 Teilnehmern** durchführen, wobei wir **8-10 kleine, mobile Kameras** in ihren Räumlichkeiten platzieren werden. Diese zeichnen das Bewegungsverhalten der Besucher auf, die mit dem Aixplorer unterwegs sind. Die Studie dauert **1-2 Stunden pro Person (bestimmt durch die Aufenthaltsdauer der Teilnehmer)**. Die so erfassten Video Daten werden in Besucherpfade umgewandelt (siehe Auswertung) und anonymisiert. **Die Video Daten werden nicht veröffentlicht.**

Vorbereitung
Es wird ein großes Schild am Eingang platziert, welches alle Besucher auf die Studie aufmerksam macht und erklärt, dass diese Daten nicht veröffentlicht werden.

Im Rahmen der Bewegungsanalyse im Rathaus, möchten wir folgende Räume mit Kameras versehen: den Ratssaal, den Roten Saal, die Werkmeisterküche, das Werkmeistergericht, den Weißen Saal, den Krönungssaal und die Eingangshalle.
Um die Studie durchführen zu können werden wir im Rathaus **8-10** Kameras anbringen.

Die Kameras
Die Kameras, die wir für die Studie einsetzen wollen sind von Typ GoPro Hero 3. Diese Kameras haben nicht nur den Vorteil, dass Sie sehr klein sind, sondern sind zudem mit Akkus bestückt und enthalten einen internen Speicher. Dadurch ist keine Verkabelung notwendig. Die Auslösung der Kameras erfolgt durch eine Fernbedienung.

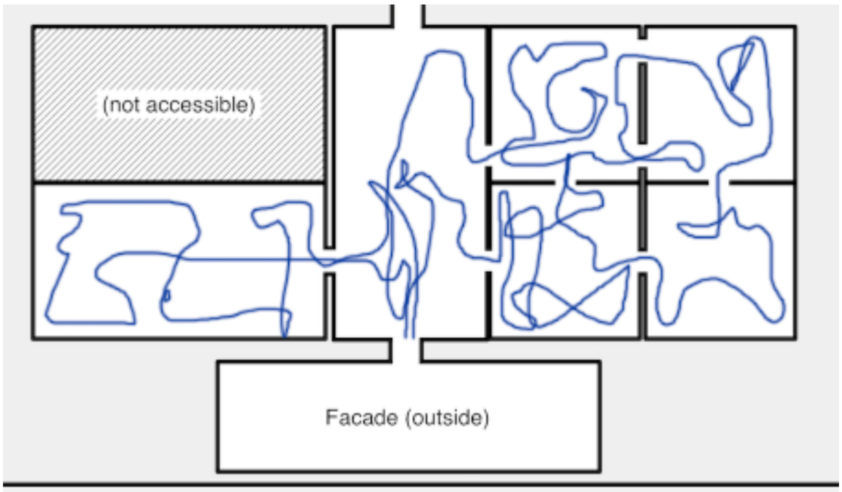
Ziel der Studie
Das Ziel der Studie ist es, zu verstehen wie sich Menschen mit Multimedia Guides in Museen bewegen. Wir streben ca. 20 Studienteilnehmer an um einen repräsentativen Datensatz erstellen zu können.

Ablauf der Studie
Im folgenden wird der Ablauf der Studie skizziert, welches gleichzeitig meine TODO Liste für die Studie ist.

1. Potentieller Teilnehmer betritt das Rathaus. Begrüßung, Vorstellung.
2. Erklärung der Studie und des Nutzen
3. Erhebung der Grunddaten. (siehe Fragebogen)
4. Ausfüllen der Einverständniserklärung
5. Ausstattung mit dem Aixplorer (kurze Einführung)
6. Auslösung der Kameras
7. Besucher bewegt sich durch das Rathaus
8. Abschluss der Studie. Als kleines Dankeschön werden Eintrittspreis und Aixplorer Nutzung von der RWTH übernommen + kleine Präsenttüte.
9. Ende

Figure A.4: Handout: Description of user study for staff and authorities (german) (page 1)

Auswertung & Veröffentlichung
Basierend auf dem Fragebogen und den Videos, werden anonymisierte Benutzerprofile (z.B. männlich, 30-40, Einzelbesucher) erstellt und Pfade der Benutzer extrahiert:



Die Pfade aller Probanden werden im weiteren Verlauf der Analyse zu sogenannten Heatmaps zusammengefasst. Heatmaps sind eine Art der Visualisierung, die erlaub häufige Aufenthaltsorte zu identifizieren. Dies kann folgendermaßen aussehen:

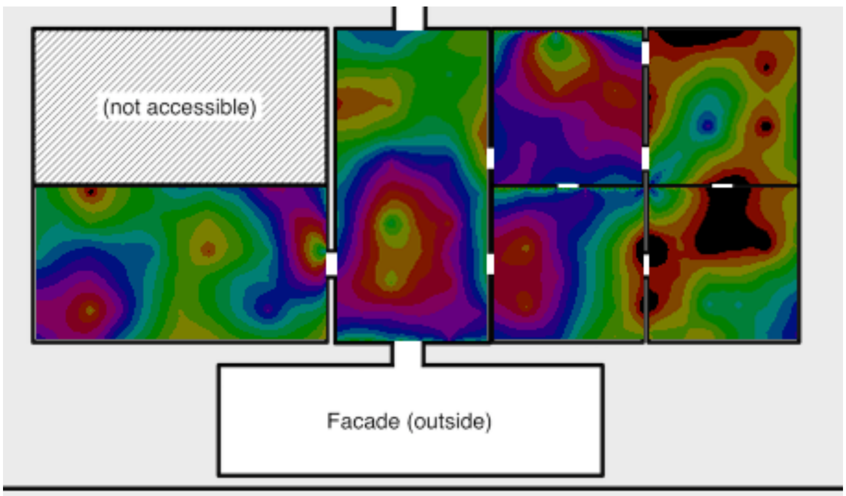


Figure A.5: Handout: Description of user study for staff and authorities (german) (page 2)

Neben den Heatmaps und einzelnen Pfaden wird ebenfalls die demographische Zusammensetzung der Studienteilnehmer veröffentlicht. z.B. "An der Studie nahmen 20 Personen im Alter von 18 - 75 teil. 43% der Teilnehmer war männlich, 57% weiblich. Über 2/3 der Teilnehmer waren in Gruppen unterwegs...".

Die initiale Veröffentlichung der Ergebnisse ist Teil meiner Masterarbeit mit dem Titel: "Movement Analysis of visitors using location aware guides in museums".

Figure A.6: Handout: Description of user study for staff and authorities (german) (page 3)

Daily Checklist

Battery	Charged?	SD Card	Videos saved?
1		1	
2		2	
3		3	
4		4	
5		5	
6		6	
7		7	
8		8	
9		9	
10		10	
11		11	
12		12	
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
R1			
R2			

Equipment

12 Cameras
24 Batteries
12 Chargers
Consent Forms
Questionaires
Poster
Documentation of Participants
Blue tack
post-its

Figure A.7: Daily checklist for study preperation

HERO3
WHITE EDITION

Smaller, lighter and Wi-Fi enabled.
GoPro's new HERO3: White Edition makes it easy to capture and share your world.

KEY CAMERA SPECS

- 1080p30 | 960p30 | 720p60 fps
- 5MP | 3 fps Burst
- Wi-Fi Built-In
- Wi-Fi Remote + App Compatible

GoPro
Be a HERO.

* Optional accessory sold separately

**Optional accessories sold separately*

Product Type:	HERO3: White Edition
Available Date:	10.21.12
UPC Code:	818279010008
Model Number:	130-01529-000
Product Name:	HERO3: White Edition
Price:	US $199.99
Included Inbox:	• HERO3: White Edition Camera • 197'/60m Waterproof Housing • Rechargeable Li-ion Battery • QR Buckle • 1 Curved + 1 Flat Adhesive Mount • Assorted Mounts and Hardware • USB Charging Cable
Single Unit: Dims	3.94in x 3.94in x 9.65in
Single Unit: Dims	10cm x 10cm x 24.5cm
Inner Pack(6): Dims	9.06in x 13in x 10.8in
Inner Pack(6): Dims	23cm x 33cm x 27.5cm
Single Unit: Weight	19.29oz/ 547g

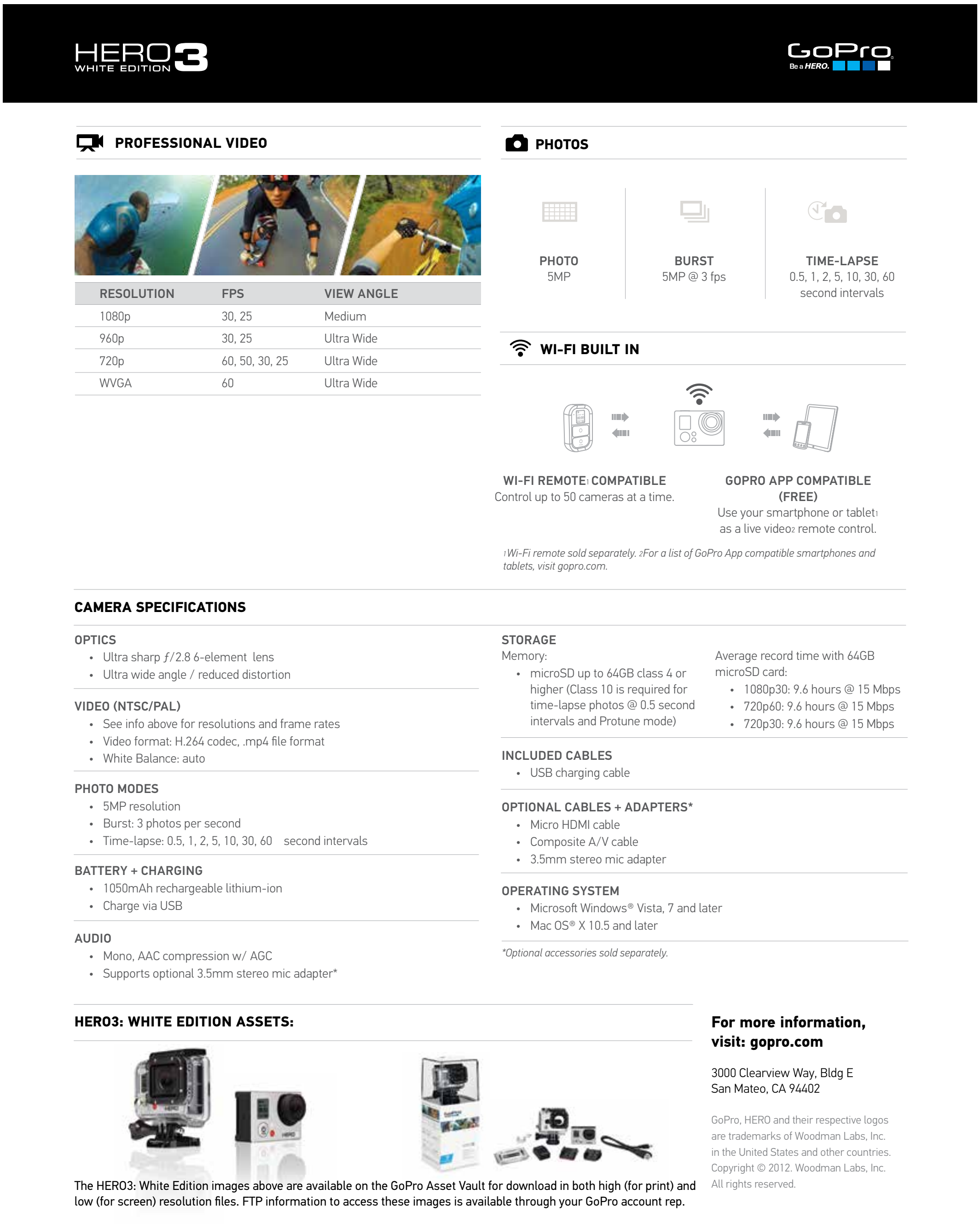
KEY BENEFITS

- Wearable, mountable design
- Immersive, wide angle capture of your favorite activities
- Professional quality HD video & 5MP photos
- Built-in Wi-Fi enables remote control via optional Wi-Fi Remote or live video preview and remote control on smartphones and tablets running the free GoPro app.
- Rugged housing is waterproof to 197ft/60M and captures sharp images above and below water
- Compatible with all GoPro mounts for attaching to gear, body, helmets, vehicles and more
- Compatible with LCD Touch BacPac™ and second generation Battery BacPac™
- Backwards compatible with older generation BacPacs™

KEY SPECS

- Professional 1080p 30 fps/960p 30fps/720p 60 fps and more video capture
- 5MP photo capture with 3 frames per second burst
- Wi-Fi Built-In
- Wi-Fi Remote Compatible (sold separately)
- GoPro App Compatible (FREE)
- 197' / 60m Waterproof Housing
- Basic mounts and hardware included for attaching to helmets, gear and more

Figure A.8: GoPro Hero 3 White Edition: Specifications (page 1)



HERO3 WHITE EDITION

GoPro Be a HERO.

PROFESSIONAL VIDEO

RESOLUTION	FPS	VIEW ANGLE
1080p	30, 25	Medium
960p	30, 25	Ultra Wide
720p	60, 50, 30, 25	Ultra Wide
WVGA	60	Ultra Wide

PHOTOS

PHOTO
5MP

BURST
5MP @ 3 fps

TIME-LAPSE
0.5, 1, 2, 5, 10, 30, 60 second intervals

WI-FI BUILT IN

WI-FI REMOTE[1] COMPATIBLE
Control up to 50 cameras at a time.

GOPRO APP COMPATIBLE (FREE)
Use your smartphone or tablet[1] as a live video[2] remote control.

[1]Wi-Fi remote sold separately. [2]For a list of GoPro App compatible smartphones and tablets, visit gopro.com.

CAMERA SPECIFICATIONS

OPTICS
- Ultra sharp ƒ/2.8 6-element lens
- Ultra wide angle / reduced distortion

VIDEO (NTSC/PAL)
- See info above for resolutions and frame rates
- Video format: H.264 codec, .mp4 file format
- White Balance: auto

PHOTO MODES
- 5MP resolution
- Burst: 3 photos per second
- Time-lapse: 0.5, 1, 2, 5, 10, 30, 60 second intervals

BATTERY + CHARGING
- 1050mAh rechargeable lithium-ion
- Charge via USB

AUDIO
- Mono, AAC compression w/ AGC
- Supports optional 3.5mm stereo mic adapter*

STORAGE
Memory:
- microSD up to 64GB class 4 or higher (Class 10 is required for time-lapse photos @ 0.5 second intervals and Protune mode)

Average record time with 64GB microSD card:
- 1080p30: 9.6 hours @ 15 Mbps
- 720p60: 9.6 hours @ 15 Mbps
- 720p30: 9.6 hours @ 15 Mbps

INCLUDED CABLES
- USB charging cable

OPTIONAL CABLES + ADAPTERS*
- Micro HDMI cable
- Composite A/V cable
- 3.5mm stereo mic adapter

OPERATING SYSTEM
- Microsoft Windows® Vista, 7 and later
- Mac OS® X 10.5 and later

**Optional accessories sold separately.*

HERO3: WHITE EDITION ASSETS:

The HERO3: White Edition images above are available on the GoPro Asset Vault for download in both high (for print) and low (for screen) resolution files. FTP information to access these images is available through your GoPro account rep.

For more information, visit: gopro.com

3000 Clearview Way, Bldg E
San Mateo, CA 94402

GoPro, HERO and their respective logos are trademarks of Woodman Labs, Inc. in the United States and other countries. Copyright © 2012. Woodman Labs, Inc. All rights reserved.

Figure A.9: GoPro Hero 3 White Edition: Specifications (page 2)

Appendix B

Software Design & Implementation

```
%{ This source code has originally been
designed by Christopher R. Wren
http://xenia.media.mit.edu/~cwren/interpolator/

I altered the code in order to have a functional
code to test my c implementation against. %}

hold off
axis([0 200 0 200 ])

% Source rectangle
X = [0; 20; 80; 200];
Y = [0; 200; 200; 0];

% point to transform
x = 46;
y = 120;

% Resulting rectangle
Xp=[0; 0; 300; 300];
Yp=[0; 300; 300; 0];

plot([X;X(1)],[Y;Y(1)],'r')
hold
plot([0 0 300 300 0], [0 300 300 0 0],'b')
axis([ -100 400 -100 400 ])

% Matrix operations
B = [ X Y ones(size(X)) zeros(4,3)
-X.*Xp -Y.*Xp ...
      zeros(4,3)        X Y ones(size(X)) -X.*Yp -Y.*Yp ];
B = reshape (B', 8 , 8 )';
D = [ Xp , Yp ];
D = reshape (D', 8 , 1 );
l = inv(B' * B) * B' * D;
A = reshape([l(1:6)' 0 0 1 ],3,3)';
C = [l(7:8)' 1];

% calculate new point from matrices
t=A*[x;y;1]/(C*[x;y;1]);

disp(t);
plot(x,y,'xr');
plot(t(1), t(2), 'ob')
```

Figure B.1: Perspective Transform Estimation Matlab Code based on Christopher R. Wren's implementation taken from Wren [1998]

Appendix C

Evaluation

Table C.1: Order of visited rooms by visitor

No	Order of visited rooms
PN1	1,6,7,6,7,6,7,6,9,6,9,6,1,5,2,3,4,5,1
PN2	1,6,7,6,9,6,1,5,2,3,4,5,1
PN3	1,6,7,6,8,9,6,1,5,2,3,4,5,4,5,1
PN4	1,6,7,8,9,6,1,5,2,3,4,5,1
PN5	1,6,9,8,7,6,9,6,7,8,7,6,1,5,4,5,4,5,4,3,2,5,4,5,1
PN6	1,6,9,8,7,6,9,6,7,6,1,5,4,3,2,1
PN7	1,6,9,6,8,9,6,9,8,7,6,9,8,7,6,1,5,4,3,2,1
PN8	1,6,7,8,9,6,7,8,9,8,7,6,8,7,6,1,5,2,3,4,5,4,5,4,5,2,3,2,1
PN9	1,5,1,6,7,8,9,6,7,8,9,8,7,8,7,6,1,5,2,3,4,5,4,3,2,1
PN10	1,6,9,7,6,1,5,4,3,4,3,2,1
PN11	1,6,9,6,1,5,4,3,2,1
PG1	1,6,7,8,9,6,9,6,7,8,7,1,5,2,3,4,5,1
PG2	1,6,9,8,9,8,7,6,7,8,7,8,6,7,8,7,6,9,8,7,1,5,2,3,4,5,1
PG3	1,6,9,8,7,6,9,8,7,8,7,8,7,6,9,6,1,5,2,3,2,3,4,3,2,1
PG4	1,2,3,4,5,1,6,9,8,7,1
PG5	1,6,7,6,9,6,7,6,1,5,2,3,4,5,1
PG6	1,6,9,6,1,5,1,5,2,3,4,5,4,5,4,5,1
PG7	1,2,3,4,5,1,6,9,7,1
PG8	1,2,3,4,5,1,6,9,8,7,6,1
PG9	1,5,4,3,2,5,4,5,1,6,7,8,9,6,7,8,7,1
PG10	1,2,5,2,3,4,5,1,6,7,8,9,6,9,6,9,6,1

Table C.2: Average Time spent by groups and single visitors

Group	Time spent in museum
Group 1	1064
Group 2	2825
Group 3	1886
Group 4	1032
Group 5	1164
Average Groups	1594
Single 1 (PG 1)	1760
Single 2 (PG 4)	711
Single 3 (PG 9)	2277
Single 4 (PG 10)	2128
Average Singles	1719

Table C.3: Time spent in tracked rooms

Time Spent in Museum by Participants (in seconds)													
	Room Number									Areas			
No	1	2	3	4	5	6	7	8	9	Staircase	Ground floor	First floor	All Areas
PN1	64	14	120	323	30	156	37	0	171	142	551	364	1057
PN2	64	44	102	171	170	171	24	0	172	146	551	367	1064
PN3	67	26	88	309	59	103	16	6	178	212	549	303	1064
PN4	176	42	104	174	163	96	8	23	173	111	659	300	1070
PN5	535	26	45	731	106	493*	218*	238*	87*	452	1443	1036	2931
PN6	538	107	128	333	215	147*	296*	393*	314*	131	1321	1150	2602
PN7	260	13	32	746	161	351*	104*	192*	332*	741	1212	979	2932
PN8	879	19	100	347	70	411*	164*	236*	274*	447	1415	1085	2947
PN9	480	140	200	379	152	82*	295*	456*	315*	134	1351	1148	2633
PN10	45	46	71	125	131	80	13	0	331	323	418	424	1165
PN11	35	45	61	120	149	76	0	0	350	327	410	426	1163
PG1	164	12	25	241	49	314	88	62	195	610	491	659	1760
PG2	60	298	344	391	161	349	230	435	242	262	1254	1256	2772
PG3	246	206	86	718	199	421	253	415	167	250	1455	1256	2961
PG4	108	61	29	36	42	62	70	31	28	244	276	191	711
PG5	236	46	206	397	292	68	36	0	279	321	1177	383	1881
PG6	113	47	195	350	297	106	0	0	279	503	1002	385	1890
PG7	113	141	99	87	131	64	15	0	106	278	571	185	1034
PG8	108	139	102	70	153	54	35	22	74	272	572	185	1029
PG9	275	171	463	530	149	62	106	53	196	272	1588	417	2277
PG10	417	6	585	164	154	52	55	58	338	299	1326	503	2128

* These values need to be regarded as PG values, since they used the guides from PG2 & PG3.

Table C.4: Demography & experience of study participants.

Demography & experience of study participants							
Participant	Gender	Age	Job	Been to the museum before?	Used Any Multimedia Guide before	Has Smartphone?	Smartphone experience
PG1	2	29	4	0	0	1	1
PG2	2	80	6	0	1	0	5
PG3	2	62	6	0	0	0	3
PG4	2	23	3	0	1	1	1
PG5	1	23	3	0	0	1	1
PG6	1	22	3	0	1	0	-
PG7	1	66	6	0	1	1	3
PG8	2	65	6	0	1	0	3
PG9	2	20	2	0	1	1	3
PG10	2	29	4	0	1	1	1
PN1	2	25	3	0	1	1	2
PN2	2	25	3	0	0	1	3
PN3	2	24	3	0	1	1	3
PN4	2	28	3	0	1	1	2
PN5	1	21	3	0	0	1	1
PN6	1	27	5	0	0	1	1
PN7	2	27	4	0	1	1	2
PN8	1	42	4	0	0	0	3
PN9	1	43	5	1	1	1	1
PN10	2	42	4	0	1	0	-
PN11	1	33	4	0	0	1	1

Job numbers refer to 1 - pupil, 2 - learning a job, 3 - student, 4 - employee, 5 - self-employed, 6 - other.
Smartphone Experience is rated from 1 - very experience down to 5 - little/no experience.
For more details, see A.1

Appendix D

CD Content (Software & Documentation)

The CD attached to this thesis contains apart from a copy of this thesis two folders.

Implementation

The folder implementation contains the source code for the Annotation Tool ″FollowThem″ and the Tracking Tool ″TrackingAnalyser″. Apart from that it contains the OpenTLD source code as XCode Project.

Material

This folder contains screenshots of the ″Tracking Analyser″ showing paths and heatmaps. Also it contains the raw files exported from ″Follow Them″.

Glossary

Aixplorer

The multimedia guide that is in place at the city hall of Aachen. It has been developed by the Media Computing Group of the RWTH Aachen. The Aixplorer uses indoor wireless tracking to obtain information about the visitor's location and supplies her with room-based information.

CORONA

CORONA is an interactive audio experience, using the Ubisense system. It is part of the coronation hall and allows the user to experience how it used to be back in the days, when the coronation hall was used for festivities. There are 5 different virtual audio sources, that the visitor can listen to, when she gets close to the source.

Ground Truth

Ground truth is a term that originates in cartography. In my thesis it referes to the actual/true position of a visitor in the room.

Peace Table / Friedenstisch

The table is a big screen, that reacts to IR sensors that are integrated into wooden blocks, that lie on the table. The wooden blocks are connected via a red line on the screen and display information about one aspect of the "Aachener Frieden", which is a historical event. The Peace Table is part of the red room (Room 5)

PNs and PGs

The abbreviation PN and PG are used throughout the thesis for *P*articipant *Normal* and *Participant* with *G*uide

Ubisense

Ubisense is a real-time location system that allows locating the user with the help of sensors. These sensors are glued to the headphones of the Aixplorer.

Bibliography

J. Berclaz, F. Fleuret, E. Turetken, and P. Fua. Multiple Object Tracking using K-Shortest Paths Optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.

Alberto Bianchi and Massimo Zancanaro. Tracking users'movements in an artistic physical space. In *The i 3 Annual Conference, Siena, Italy*, volume October 20-22, pages 103–106, 1999.

S. Calderara, R. Cucchiara, and A. Prati. Bayesian-competitive consistent labeling for people surveillance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2):354 –360, feb. 2008. ISSN 0162-8828. doi: 10.1109/TPAMI.2007.70814.

Luca Chittaro and Lucio Ieronutti. A visual tool for tracing users' behavior in virtual environments. In *Proceedings of the working conference on Advanced visual interfaces*, AVI '04, pages 40–47, New York, NY, USA, 2004. ACM. ISBN 1-58113-867-9. doi: 10.1145/989863.989868. URL `http://doi.acm.org/10.1145/989863.989868`.

A. Criminisi, I. Reid, and A. Zisserman. A plane measuring device. *Image and Vision Computing*, 17(8):625 – 634, 1999. ISSN 0262-8856. doi: 10.1016/S0262-8856(98)00183-8. URL `http://www.sciencedirect.com/science/article/pii/S0262885698001838`.

Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01*, CVPR '05, pages 886–893, Washington, DC, USA, 2005. IEEE Computer Society. ISBN 0-7695-2372-2. doi:

10.1109/CVPR.2005.177. URL http://dx.doi.org/10.1109/CVPR.2005.177.

F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua. Multi-Camera People Tracking with a Probabilistic Occupancy Map. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):267–282, February 2008.

Gero Herkenrath and Jan Borchers. Aixplorer - mobile audio and multimedia tour guide. URL http://www.aixplorer.de/.

Z Kalal, J Matas, and K Mikolajczyk. P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints. *Conference on Computer Vision and Pattern Recognition*, 2010.

Daniela Petrelli, Antonella De Angeli, and Gregorio Convertino. A user-centered approach to user modeling. In *Proceedings of the seventh international conference on User modeling*, UM '99, pages 255–264, Secaucus, NJ, USA, 1999. Springer-Verlag New York, Inc. ISBN 3-211-83151-7. URL http://dl.acm.org/citation.cfm?id=317328.317366.

Kingkarn Sookhanaphibarn and Ruck Thawonmas. A movement data analysis and synthesis tool for museum visitors' behaviors. In *Proceedings of the 10th Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing*, PCM '09, pages 144–154, Berlin, Heidelberg, 2009. Springer-Verlag. ISBN 978-3-642-10466-4. doi: 10.1007/978-3-642-10467-1_12. URL http://dx.doi.org/10.1007/978-3-642-10467-1_12.

Eliséo Véron, Martine Levasseur, de gestion et de marketing Société d'organisation, and Bibliothèque publique d'information (Paris). *Ethnographie de l'exposition : l'espace, le corps et le sens*. Centre Georges Pompidou, Paris, 1983. ISBN 2-902706-01-4.

Christopher R. Wren. Perspective transform estimation, 12 1998. URL http://xenia.media.mit.edu/~cwren/interpolator/.

Massimo Zancanaro, Tsvi Kuflik, Zvi Boger, Dina Goren-Bar, and Dan Goldwasser. Analyzing museum visitors' behavior patterns. In *Proceedings of the 11th* *international conference on User Modeling*, UM '07, pages 238–246, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN 978-3-540-73077-4. doi: 10.1007/978-3-540-73078-1_27. URL http://dx.doi.org/10.1007/978-3-540-73078-1_27.

Index

Typeset March 28, 2013