

Virtual Mirror Gaming in Libraries

Marijn Speelman and Ben Kröse
Informatics Institute, University of Amsterdam
Kruislaan 403, 1098 SJ Amsterdam
contact: b.j.a.krose@uva.nl

Abstract

This paper presents a study on a natural interface game in the context of a library. We developed a camera-based Virtual Mirror (VM) game, in which the player can see himself on the screen as if he looks at a mirror image. We present an overview of the different aspects of VM games and technologies that can be used. For this study a novel framework and prototype game VMQuiz was developed. The prototype was used to evaluate the difference in experienced enjoyment between gaming via our camera interface, and a traditional mouse based interface. The results from the evaluation show that children like playing games in the library and like book related games. The VM game was not experienced to be more enjoyable than the mouse version.

1. Introduction

Over the past years, gaming has become one of the main activities of children and young adults. Using desktop PCs, consoles like the Playstation and Nintendo Wii, and more recently mobile phones, kids spend a great deal of time playing games, alone, over the Internet or with friends in real life. A study done by the NICAM¹ in 2003 shows that almost all kids play a game once in a while and two-third play games on a daily basis. With our life increasingly taking place in the digital world, the popularity of reading books and consequently library visits by children is decreasing.

Games and books have many similarities that make games interesting for libraries. Games can be informative, or entertaining, like books. We were interested whether playing games inside a library would encourage children to visit the library more often, and keep them in contact with the 'old' medium.

Playing games in (semi)-public places puts some constraints on the interfacing. Traditionally gaming is done using a gaming console using some sort of joystick or a desk-

top PC using a mouse and keyboard. New ways of interaction emerged over the years and have reached the consumer market in great numbers. The Nintendo Wii [11] is the latest example of a successful product that has a non-traditional interface. The Wii uses sensors to detect the orientation of the controller. Control can also be done using computer vision techniques like motion detection. Such games can be called webcam or camera-based games. If the player sees himself on the screen, it is called a Virtual Mirror (VM) game [31]. Sony has developed the EyeToy [7], which is a small USB webcam that can be attached to the Playstation 2 to play these kinds of games.

In particular webcam games could be an interesting addition to libraries. They are distinctly different from traditional games because they are relatively new and unknown, provide a new way of interaction and are not available for every child in their home.

Following these thoughts, the goal of our research is: 1) To compare the level of enjoyment between traditional gaming and VM gaming and 2) to determine if the library is a suitable environment to deploy such games. To test these hypothesis we created a prototype game named *VMQuiz*. A software framework was developed for this game and to use in future studies on webcam gaming. It supports rapid development of multimedia applications and implementation of existing computer vision software libraries.

In Section 2 we describe how gaming fits the context of a library. In Section 3 we describe how computer vision can be used in gaming and which technologies and approaches of movement analysis can be used. In Section 4 we describe the technical and functional aspects of the prototype that has been created. Finally, Section 5 presents the experiment we conducted to evaluate the prototype and the results of these experiments and Section 6 the conclusion.

2. Gaming in Libraries

Libraries are not educational institutions like schools, but they do have a great value in teaching children certain values, skills and increasing general knowledge. One of the main goals for libraries is of course to increase literacy for

¹Nederlands Instituut voor de Classificatie van Audiovisuele Media

children. Games can be a great addition to books. In a certain way they are ‘just another’ medium next to books. One of the most clear similarities is that both books and games often try to tell a story. The writer tries to let the reader feel as if he takes really takes part of what is happening in the story. Story lines and genres in books and games are often closely related. For example there is the game Warcraft by Blizzard which is inspired by The Lord of the Rings by J.R.R Tolkien. The story line of Warcraft has even been expanded in a book series. This shows games and books can be closely related. Apart from their links in genre and story telling, playing games could increase skills like reaction time, resource management, team work, critical thinking, making quick decisions, fast reading and coping with stress. It can also increase knowledge on subjects like historical times [26].

Libraries acknowledge the value of games and have been lending out games to customers for many years. The last few years libraries have been more actively using games inside the library. An overview of the intersection between gaming and libraries is given in [23]. Many libraries have PCs for kids with pre-installed games. Often these games have some educational value, like teaching basic math skills using a fun story and interesting setting that relates to the mind-set of children. Next to their educational purpose, libraries hope games make a stay at the library more fun. Although libraries maybe do not have that ‘dusty’ reputation it had in the past, there is still much work to be done. There are large groups of children that do not visit the library at all.

In the United States there have been several libraries that held LAN parties on weekends or that have been trying to put innovative gaming applications on the library floor. In the Netherlands there have also been libraries, like the library of Vlissingen in the province of Zeeland, that organised gaming days [9]. There are several websites in the form of communities with blogs and wikis where librarians share their experience on gaming activities like LAN parties² [1]. An important aspect of these gaming events is the social interaction between the players. This effect can be even stronger with games with a physical interface, like the Wii or Virtual Mirror games.

3. Computer Vision in Gaming

A new way of controlling a game is to use camera input (often a webcam) and computer vision techniques. The computer program tries to understand the contents of a scene by determining specific features like faces or by detecting motion in a stream of images. The result of this analysis can be used to create a game that is based on phys-

²A Local Area Network (LAN) party is a gathering where several gamers play games against each other

ical movement. In the following sections we describe such a system and what kind of techniques for analysis can be used.

3.1. Virtual Mirror Effect

A Virtual Mirror (VM) game is a type of webcam game where the player is being recorded using a digital camera and sees himself in real time on the screen or projected image in front of him. This creates the illusion of looking at a mirror. The result is a combination of the camera image and (virtual) game graphics. A graphical representation of a physical setup of such a system can be seen in figure 1.

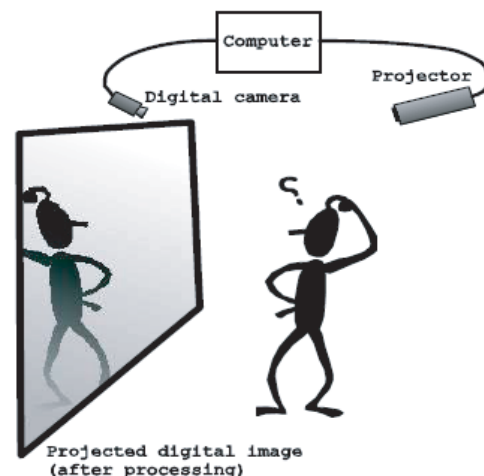


Figure 1. Virtual Mirror setup, from [31].

One of the first descriptions and applications of a VM game was VIDEOPLACE (1985) [21]. VIDEOPLACE was designed to be an addition to traditional telephony so both players could see each others silhouette, but it was also used to play a little game. A creature called CRITTER climbs the silhouette of the player, slides down on his arm and repeats this.

The mirror effect comes with some difficulties. Because the image is recorded from a static location, the eyes of the user do not follow the image as with a real mirror [31]. A problem related to this is eye contact. With a real mirror the person keeps eye contact as he moves in front of the mirror. There are some technical and practical solutions for this, but those are not really needed for intuitive interaction. Especially when playing a game the player will not really notice where his eyes are looking at because he is most likely looking at his hands or game graphics on the screen.

The game graphics can be virtual objects that can be manipulated. The objects are computer generated and react to the movements of the player. There are different types of objects:

- Static objects. This is the static game interface of the

game, like the current score of the game.

- Buttons. These objects can be activated or touched by a player. By generating activity by moving at the location of the button, it is being pushed.
- Movable objects. These objects react to the movement of the player. By analysing the player's movement the player can drag the images across the screen.

In the next section we will discuss how interaction with the aforementioned buttons and movable objects can be realised.

3.2. Technologies for VM Games

There are several techniques that can be used to detect and analyse player movement and player features. In this section we describe the techniques that are most relevant for VM games. First, there is simple motion detection that uses image differencing to determine moving areas. Secondly, there is face detection to detect a human face and thus the presence of a player. Finally, there is object detection, which is the detection and tracking of predefined objects using shape or colour.

Other frequently used techniques are background subtraction and skin colour segmentation, which are beyond the scope of this paper.

3.2.1 Motion detection

Motion detection or motion sensing uses simple image differencing. The RGB colours are compared between two subsequent frames. When the difference between the RGB value of two corresponding pixels is greater than a certain threshold, the game detects movement. These single changed pixels can then be grouped and used to determine larger areas of movement. This type of motion detection can be used to touch objects in a game, like hitting a ball with the head or foot (Virtua Soccer [8]) or touching a virtual button. Because the game does not know the direction of the motion with this method, it can not be used to push around movable objects.

To track the direction and speed of motion, optical flow can be used [13]. For each frame the displacement compared to the previous frame is calculated for certain areas. These areas can be the virtual objects that are being tracked. It is not needed to calculate the optical flow for the whole image. A comparison between motion detection and optical flow is done in [31]. The study shows using optical flow for touching virtual buttons performance faster and more accurate than motion detection, especially for complex menus.

Often when analysing body movement with a camera there is no distinction made between different limbs like arms or legs, just the relative frame by frame difference is

used. One can look at the movement of the limbs from the front of the player, or at the movement of the whole body in space from the top. An example of a combination of both is the experiment in [22] where a wide-angle lens looking down onto the playing field was used to keep track of spatial orientation in the map of the game. Also the movement of the arms was used to execute different game commands.

3.2.2 Face detection and -tracking

Face detection should not be confused with face recognition. *Detection* only detects if and where a face is present in the image, while *recognition* tries to determine who's face it is. Face tracking is done by detecting the face in each subsequent frame.

There are various methods for the detection of faces in images or sequences of images. An extensive survey of current face detection methods is given in [30]. In our prototype game (section 5.1), we used the face detector proposed by [28] to detect frontal faces. It is scale-invariant, meaning all sizes of faces can be detected in the image.

The added value of face detection in gaming was studied by [29]. It concluded that the use of face detection has several advantages compared to traditional motion detection. The main two reasons are:

- Increase of the ability of the game to detect if a person is present and the location of that person.
- Improvement of the game experience using role playing by adding virtual objects on the player's body, for instance a hat or jewels around a player's neck.

Other advantages are that the player does not have to move to let the game detect its presence, the player can be recognised without doing anything. This could be used to automatically start the game when a player enters the game area. With face detection it is also possible to detect the number of separate players so each player can be given a separate role in the game. Disadvantages are that face detection is more reliant on external factors like skin colour and lighting. The calculation time is longer and implementation is more difficult than when using motion detection.

Face tracking and object tracking are closely related. A face is actually just an object with certain detectable features, like skin colour and the area of the eyebrows, eyes, nose and mouth. Faces are more dynamic than most objects because they can change, for instance when someone laughs.

3.2.3 Object tracking

Real objects can be used to interact with the virtual world of a game. An example is using a real paint brush to paint a virtual painting [6]. Giving the object a distinct colour

(like bright green) makes it easy to locate and track. Object tracking is less dependent on the background than motion detection, but can be more sensitive to changing light conditions. A comparison between motion detection and object tracking is done in [17]. The study concludes that, especially using many virtual buttons, object tracking performs faster and more accurate. A disadvantage of using objects is the need to constantly hold the object. Another possible reason to not use objects, is that the player only has to move his arms and not his whole body.

3.3. Existing webcam & Virtual Mirror games

There are several existing webcam games. The oldest camera-based game is considered to be VIDEOPLACE, discussed in section 3.1. Other applications are the Vivid Group Gesture Xtreme System (1986) [4], the Reality Fusion GameCam (1999) [5], the Sony EyeToy (2004) for the PlayStation2 [7] and Microsoft Xbox Live Vision (2006) [10]. Motion detection and face recognition are the most used computer vision technologies in these games. For example the Xbox Live Vision software supports capturing the face of the player and using the face on a virtual character in the game. Over the past few years, webcam producers like Logitech and Philips have also been implementing face detection and tracking in their webcam software and including small games based on gestures. This shows these technologies are becoming more common for consumers.

There are also small Flash games [2] that can be played instantly inside a browser and only require the Adobe Flash plugin and a webcam. No additional hardware or software is needed, unlike the EyeToy and Xbox Live Vision camera. Most of these games use only primitive image differencing motion detection.

4. Game evaluation

Evaluating games can be a complicated task. For game developers it is increasingly important to evaluate the games during development and get constructive feedback to deliver enjoying and well selling games. Like in this study, game evaluation can also be a research tool to compare certain aspects or features of games.

There are many terms that are related to the experience of playing a game, for example: enjoyment, usability, fun, playability and flow. In more general terms, *fun* is the most important aspect of whether someone likes to play the game or not. If a game is not considered to be fun it is simply not being played, or at least not for long. Another common term is usability, which is defined by the ISO 9241-11 standard as a combination of three measures: effectiveness, efficiency and satisfaction. A brief overview of the link between usability and games is given in [20].

Different tools can be used to evaluate games. Cam-

eras are useful to record the behaviour of the player. The video can later be used to analyse movements and to detect emotions like frustration and confusion. It is also possible to gather in-game statistics. For example the score, playing time and game actions of a player can be used to make assumptions about the difficulty and enjoyment of the game. Interviews can be used to gather qualitative information about the game or get statistical data. Questionnaires can be used to gather quantitative data about the game.

A game can be evaluated using expert analysis or user studies. Heuristics are often used with both. These heuristics are based on theoretical and empirical research. One of the first empirical studies we found on the use of heuristics and game evaluation is [24]. Different versions of a dart game were compared during a user study. The total time playing the game was taken as a indication of the enjoyability. From these results heuristics were extracted. Over the years several other studies on heuristics have been done. See [18] for an extensive overview of usability heuristics literature in gaming. Other studies are [16] in which the measurement of playability with heuristics is discussed, and GameFlow [27].

The GameFlow Model heuristics are derived from literature (i.e. [18] [16]) and is combined with the *flow* theory [15]. The idea of flow is that a person is fully involved in the experience and completely focussed on the task it is doing. When a player has this experience, the enjoyment of a game is likely to be greater and the game will be played longer. A short and clear overview of flow and gaming is given in [14]. In the GameFlow model eight criteria are given to measure the enjoyment of a game: concentration, challenge, skills, control, clear goals, feedback, immersion and social interaction. We based our questionnaire the criteria 'control', 'skills' and 'immersion' because only the criteria that are related to the interface and control were used. Furthermore we added questions to measure enjoyment directly, from a study on the enjoyment of interaction [19] and motivation [25].

The study done in [12] is comparable to our study in the sense that the difference between two game interfaces is researched. They developed an outdoor Augmented Reality (AR) Space Invaders 3D game and compare that with a PC version of the game. A questionnaire with 19 questions and likert-scale answers is used to measure the satisfaction. The AR version appears to be more satisfying than the PC version. When comparing two versions, a between-subjects or a within-subjects design can be used. This means either of the versions can be played by each participant or each participant can play both versions. The advantage of a between-subjects design is that the order of the played versions and previous version does not influence the result. We used a between-subjects design because of this. The disadvantage is that it is not possible to compare the opinion of

each player for both versions.

5. VMQuiz, a prototype VM game

A VM prototype game was created to compare the experience with a traditional mouse interface and a physical interface that uses the webcam. The game can be played with either of the input devices. A framework using a combination of Adobe Flash and C++ was developed to be able to quickly create games that include animations, sound and vector graphics, while at the same time maintaining the opportunity to use advanced image analysis.

The game was developed in cooperation with the Bibliotheek van Almere, the main library of the city Almere, Netherlands. Beforehand a list of important aspects and requirements of the game was put together. These requirements originated from meetings with the library, existing game heuristics from literature and personal gaming experience.

In the next section (5.1), the game itself is described, as well as some design choices. In section 5.2 the technical implementation of the game is discussed. Finally, some library specific aspects are discussed in section 5.3.

5.1. Description of the game

We have chosen to create a book quiz that asks questions about popular books for children. Children are familiar with quizzes and the goal of the game can be easily understood: giving the correct answers to the questions. Questions were related to well known books like Harry Potter, Pippi Langkous, De Kameleon, and (Dutch) book authors like Paul Biegel and Carry Slee.

The game was developed for boys and girls from the age of 7 to 11 and the questions were adjusted to that age group. There are two types of questions, *Multiple Choice (MC) questions* and *Sort questions*. The game consists of 12 MC questions and 3 sort questions. Most of the questions were created by the library of Almere. The number of MC questions is greater than the number of sort questions, because they are faster to answer. The questions were mixed in advance to give some variation. With the MC questions, the player has to pick one of the four possible answers. An answer is chosen by sliding it to the side of the screen, as shown in figure 2 (webcam input). In figure 3 a screenshot is shown of the version that uses mouse input. The Sort questions ask the player to match 4 images and words. The images have to be dragged to the designated drop zone beneath the corresponding words. The images originate from books. An example can be seen in figure 4.

The webcam version requires the player to move across the screen. Often ducking or crawling is needed to give an answer. The player stands in front of the webcam and sees himself on the projected image on the screen. The game



Figure 2. Webcam: Answering a Multiple Choice question by sliding the correct answer to the corner.

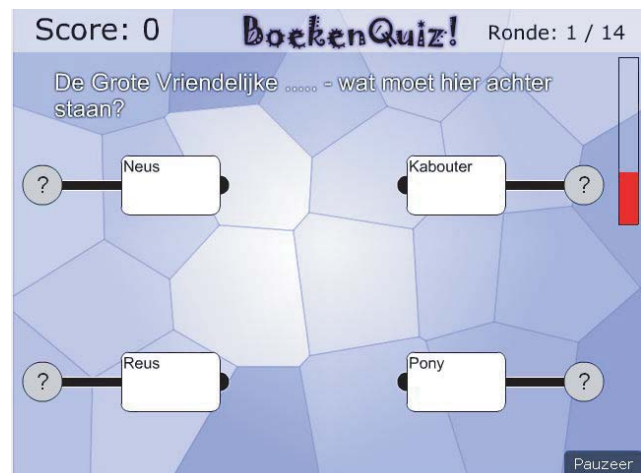


Figure 3. Mouse: Screenshot of a Multiple Choice question played using the mouse.

graphics are added to the webcam image. The mouse version works by just clicking or dragging the objects with the mouse.

The game is aimed to take about 5 to 10 minutes to play. That should be enough to give the child a good idea of what the game is like and at the same time will not put too much stress on the child. For each question a bar is displayed in the top right corner to show the remaining time of the current question. When the time is up, the player can still give the right answer but receives only half of the points he would normally get. For the MC questions the player has two tries to get it right. Giving the correct answer in the first attempt within the time limit gives 100 points, a correct answer the second try gives 75 points. When two wrong answers are given the correct answer is shown and the game continues with the next round. The sort questions give 100 points. When an image matches the word it is hovering at



Figure 4. Child answering a sort question by pushing the images (the two dogs, horse and head) across the screen.

that moment, the image locks to the word and can not be moved anymore. For example, the bottom right image in figure 4 is locked to the white coloured dropping zone.

The progress of the game is shown in the top of the screen. The current round and total number of rounds gives an idea of how long the game will be. The current score is given to show the progress of the player and to increase the motivation to get a higher score.

For most questions a small sound sample is played that relates to the current question. Examples are the sound of a howling wolf with horror questions and the Harry Potter tune with a question related to Harry Potter. Giving the wrong or correct answer plays respectively a buzzer or trumpet sound to give clear feedback of the players actions. When all questions are answered the game is finished and the final score is displayed on screen.

5.2. Technical implementation

A framework was created to allow rapid development and expansion of the current prototype. The framework consists of a front-end game client (Adobe ActionScript 3) and a server (C++) that analyses the image stream from the webcam. The analysis is done using the OpenCV computer vision library [3]. See figure 5 for a graphical representation of the application components. To minimise network delay, all applications are run on the same computer.

The client takes care of the game graphics, sounds and actual gameplay. For the client the Adobe ActionScript language was chosen because it is a high level scripting language with support for many multimedia formats and is very suitable for game creation, animation and vector graphics. The latest version, ActionScript3 was used because it is faster and has better functionality for accessing raw image and network data than previous ActionScript

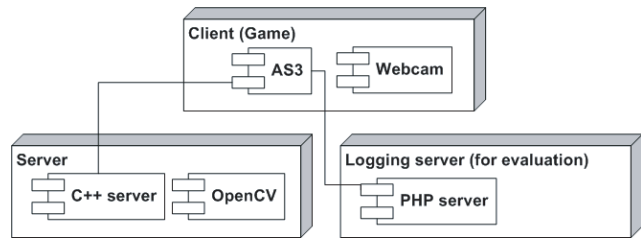


Figure 5. Application components of the framework.

versions. ActionScript code is compiled into a executable Flash file (.swf) that can be run in the browser. The server has to be started separately. Additionally a simple PHP³-script was used to log the game progress and scores to give insight in the amount of time that is needed to complete certain questions and the game as a whole.

The webcam input is read by the Flash client and sent to the server for analysis. To have smooth movement feedback, it is necessary to send image data very frequently. For face detection the image data is sent every 50ms. For the optical flow calculation the data is sent every 30ms. On a computer with a 2.8Ghz Pentium 4 processor, the average time needed for face detection is 20ms. The optical flow takes about 8ms to calculate. On top of this, there is an overhead of 10-15ms for transferring the content between the applications.

Each frame is sent in a package that contains the following data (total of 77609 bytes):

Mode A number representing the mode of the game, which can be *detect face*, *optical flow* or *inactive*.

Image Data A scaled down bitmap of the current frame of the webcam stream. The original webcam size is 640x480 pixels which is re-sized to 160x120 before sending to speed up both the sending of the data as well as the analysis of the image.

Number of Objects to Track A number which tell how many virtual objects there need to be tracked.

Object locations The x and y coordinates of the virtual objects being tracked.

Just a black and white image is needed for face detection and optical flow, so the colour channels are not used. We used the face detector that is available in from the OpenCV library, as discussed earlier in section 3.2.2. The face detection is used to automatically start the game when a players' face enters the playing area. The movement of objects is done using optical flow as discussed in section 3.2.1. After analysis the result is sent back to the client. For face detection the x and y coordinates, width and height of the

³Website: <http://www.php.net>

detected face are transmitted. For optical flow the *new* locations of the virtual tracked objects are sent back to the game client and processed further.

The mouse game is the exact same as the webcam game, with the single difference that the mouse game has the image analysis turned off and instead shows a background image and not the webcam stream.

5.3. Playing the game inside a library

When playing a game in an environment like a library, there are things that have to be taken into account:

- **Starting the game** The game should be played with as little mouse or keyboard activity as possible. For instance it should be able to start automatically when a player enters the game area. This can be done using face detection or simple motion detection. Face detection is more accurate because then a passing person will not trigger the game by accident.
- **Background** The game should be set up in a way that background movement is minimised. A room divider can be used to shield off the game area. Background subtraction and face or body recognition could be used to filter out background noise. The background colours, background pattern and lighting conditions can greatly influence the performance of the image analysis.
- **Position and distance of the camera** Different setups are possible. The player must be able to reach the corners of the screen and must be able to duck or crawl while staying in sight of the camera. The height of the player also affects the ideal distance and panning of the camera.
- **Movement space** The player should have plenty of room to move around, without accidentally touching objects or other people.
- **Sound** When playing a traditional game it is easy to use a headphone or small sound boxes so other library visitors will not be interrupted. Children playing a webcam game are further away from the screen and computer which could result in a more noisy environment.

An example setup can be seen in figure 6. Using a curved wall the sound will be somewhat contained in the gaming area and there will be minimal background noise of spectators or people walking by. The cross is where the player stands, watching to the screen in front of him where the image is projected on. Aiming the webcam at eye level will improve the mirror effect. The beamer can be attached to the ceiling or placed in front of the player. The webcam can be positioned on different locations:

- A. Beamer and camera on the table in front of the player.
- B. Camera to the left or right of the projected image.
- C. Camera recording at eye level through a little hole in the screen.
- D. Camera behind a semi-transparent screen.

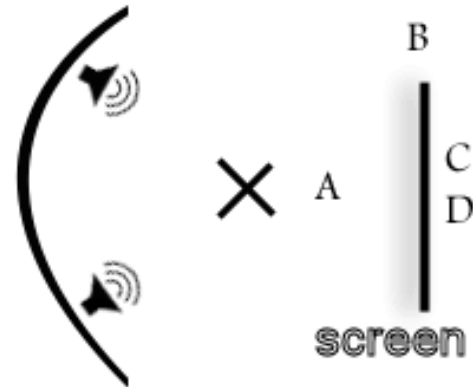


Figure 6. Example of a webcam game setup.

6. User study

The goal of our user study is to determine the level of enjoyment for both versions of the game. The second goal is to gain experience with these kind of webcam games and to find out if children enjoy playing games that relate to the books they read.

Our study researches the following hypotheses:

1. The webcam interface version is experienced to give more enjoyment than the traditional mouse interface version.
2. Children like it when they play a game that relates to the library context they are playing in.

6.1. Participants

The experiment was executed during two days at the Bibliotheek Almere, the Netherlands in December 2007. In a relatively quiet corner of the library we set up a beamer, projection screen and laptop with integrated webcam. To minimise background noise the area was partially excluded by a room divider, as seen in the background of figure 4. In figure 7 a picture of the evaluation setup is shown.

We recruited 60 children as participants of this study. Half of them, 30 children, played the webcam version; the other half played the mouse version. This results in a between subjects setup. One-by-one the participants were put in into the two different categories. Recruitment was done



Figure 7. Child evaluating the game inside the Bibliotheek Almere.

using convenience sampling in the library and an announcement on the website of the library. This announcement was also copied by the local newspaper.

The group of participants consisted of 27 boys and 33 girls. The mean age was 9,9 with a standard deviation of 1,7. The participants did not receive a reward for their participation.

6.2. Procedure

Each participant was asked to play one of the versions of the game. First a short explanation was given. Then it was explained the game was about well known books. After that the children were told how to control the game. Depending on the height of the player the camera angle was adjusted and the player positioned in the centre of the screen.

Then the game was played until it was completed. After the completion of the game the participant was asked to answer a few questions about the game. When there was unclear behaviour of the game or the participant became visibly frustrated, additional instructions were given. The mouse version took around 5 minutes, the Webcam version took 5 to 10 minutes to play. The answering of the questions took about 5 minutes for each participant.

6.3. Instruments

After playing the game, the participants were interviewed. Based on the literature study in section 4 we constructed a questionnaire. To measure the enjoyment we used relevant concepts from the GameFlow heuristics, namely Control, Immersion, Skills, with the general concept Enjoyment. The result is the list of 15 statements as showed in Table 1. To test the second hypothesis, 3 questions that are related to the content and context of the game, respectively books and the library are added. Answers were based on a 3 point likert-scale. The questions were posed in random

Table 1. The questionnaire used in the experiment

Control	
C1	I found it difficult to control the game. *
C2	I found it easy to pick the right answer.
C3	I understood quickly how to control.
Immersion	
I1	Some moments I was really concentrated on the task.
I2	I had the idea that I was completely into the game.
I3	Some moments I forgot that I was in the library
Skills	
S1	I played with a Ninetendo Wii before.
S2	I play computer games.
Enjoyment	
E1	I enjoy playing the game.
E2	I find the game enjoyable.
E3	I find the game fascinating.
E4	I find the game boring. *
E5	I am glad that the game is over. *
E6	I would like to come back and play the game again.
E7	I experienced the game as schoolwork. *
Library specific	
L1	Did you know the books the questions were referring to?
L2	Did you like the fact the questions were related to books?
L3	Would you visit the library more often when you could play this game and other similar games?

* Contra indicative

order. Both versions recorded the game progression, like the time of the whole game and each separate question. The game scores of the player were written on the questionnaire form.

After these questions, the children were asked three open questions about what they liked and did not like about the game, and if they knew of something that could improve the game. Finally the background variables age and sex were collected.

6.4. Results

All children in the test completed the game. Some children had difficulties to complete the sorting part of the quiz. With some help of bystanders they were able to continue to the next round. Overall the group of children was quite experienced with playing computer games. More than half of the participants had experience with the Wii and played games more than 4 times per week. The average score of the players was 760 with a standard deviation of 200.

We executed a reliability analysis on the groups of question to assess if they measure the same concept. For this the

Table 2. Mann-Whitney U test comparing the two different interfaces

Code	Mean webcam	Mean mouse	Mann-Whitney U	Asymp. Sig. (2-tailed)
C1	2.07	1.17	182.500	.000*
C2	1.87	1.50	320.000	.034*
C3	1.50	1.13	324.000	.016*
I1	1.23	1.23	438.500	.812
I2	1.53	1.87	327.000	.047*
I3	2.40	2.07	347.500	.097
E1	1.07	1.00	420.000	.154
E2	1.13	1.07	434.000	.621
E3	1.47	1.27	404.000	.397
E4	1.03	1.00	435.000	.317
E5	1.53	1.43	432.000	.759
E6	1.37	1.33	415.000	.504
E7	1.47	1.67	380.000	.241
L1	1.80	1.90	412.500	.482
L2	1.20	1.30	406.500	.369
L3	1.30	1.23	445.000	.926

* Significant at the 0.05 level (2-tailed).

Cronbach's alpha statistic was used. For this test we used the acceptable value of an alpha higher than 0.7. The Enjoyment concept had a Cronbach's alpha of .719 after deletion of question E7. None of the other concepts showed a sufficiently high alpha, which means no questions could be grouped together.

Correlations between questions and variables were calculated using Spearman's rho. What clearly stands out is that there is a strong negative correlation ($\rho = -0.773$, $p = 0.000$) between the played version and C1. This means the webcam interface was experienced to be far more difficult to control than the traditional mouse interface. We tested this with a nonparametric Mann Whitney test, the results are shown in table 2

For research hypothesis 1 a nonparametric Mann Whitney test was done with the two 30 participant groups of players. Question E1 to E7 were used for this. We also studied the Immersion concept and the library related questions. A score of 1 means 'agree' for positive questions and 'not agree' for negative questions. The result is shown in table 2. For the Control concept the results correspond to the results of the Spearman's test. It is significant that the webcam version was found more difficult to control. The mean score for the enjoyment questions was slightly higher in the webcam condition than in the mouse condition, but the difference is not significant.

For research hypothesis 2 the mean score was calculated and compared from the last three questions (L1, L2, L3). As shown in table 2 there was no significant difference between the two interfaces. The children really like the fact that the questions were related to books. They also believe they will

visit the library more often if they could play such games. Not surprisingly, there is a positive correlation between age and L1, the familiarity with the books the questions were referring to. In general these three questions show the children have a positive attitude towards library or book related games.

6.5. Discussion

The goal of this experiment was to measure the enjoyment of both interfaces. A higher score on the questionnaire should result in a better and more enjoyable game. In general there were no significant differences between the two interfaces. This suggests that both interfaces have the same perceived enjoyment. Another explanation is that the questionnaire is not entirely suited for correctly measuring enjoyment and flow. It is likely that there are not enough questions to get an accurate measurement of all aspects. The results of the questions appear to be highly influenced by the experienced difficulty to control the game.

There were several children that played both versions of the game. The interview was done before playing the other version to not influence the results. All children that played both versions said they liked the webcam interface more, because they liked the movement aspect of it. They were visibly enthusiastic when playing the game, with a few exceptions. One participant got quite frustrated with a sort question and almost gave up. He told us he was used to the capabilities of the EyeToy, which allows faster and less accurate movements.

The questionnaire results show that the children liked the idea of playing book related games. This also shows from the collected comments. Many children said they enjoyed the game because it asked questions about books they know. Some children also said they learnt some new facts about books. They even expected to visit the library more often when these kinds of games can be played. They were also interested in playing our game again.

The played game was a prototype that had not been played by children before. It had some short-comings that negatively influenced the enjoyment of the game. For example the controls for the sorting questions were found too difficult for a number of children. Further improvement and a more robust movement analysis is needed to present a fair 'opponent' to the more easily controlled mouse interface. Looking at the scores and comments of the children, the difficulty of the game questions seems to be right. The age of the child shows a correlation between how much the children know about the books, but this did not affect their score enough to be significant.

7. Conclusion & Future work

In this paper we presented an overview of the current state of webcam games and the techniques that can be used to analyse the webcam images. A literature study was done on the evaluation of games, which showed there are multiple studies that use heuristics for evaluation but none really fit webcam games well. We showed how gaming and in particular webcam games can be played in libraries and what the possible benefits are, like more frequent library visits and making visits more fun for children.

Based on the results of the user study using our VMQuiz prototype game, we can conclude that there was no significant difference between the measured enjoyment of the player playing the traditional mouse game, and the one with the webcam interface. Both versions scored high. After the experiment children were allowed to play both conditions. From the comments we collected during the interviews, the children appeared to enjoy playing webcam games and thought it was a fun way to interact with the game.

The second research objective was to test the hypothesis that children like playing games that relate to books and that they like playing them in the library. From the results of the questionnaire it can be concluded that this hypothesis is valid. The children were also visibly enthusiastic when playing both versions of the game and predicted they will visit the library more often when these kind of games can be played in the library.

Finally, a novel software framework was developed that aims to make development of webcam games easier and quicker and make use of more advanced computer vision algorithms. It consists of a Flash client for the actual game and media like sound and graphics, and a C++ server application that is able to use existing image analysis algorithms.

The research has raised a number of issues with the current prototype and questionnaire. To further analyse the effect of the game interface, the prototype should be improved. Movement analysis can be made more robust by improving the current methods and using more accurate movement tracking algorithms. Also different types of questions and game objectives can make the game more easy and fun to play. The questionnaire can be expanded with more questions about the specific interface. Questions that ask more directly about the enjoyment could also be added.

8. Acknowledgements

Thanks to Zoran Zivkovic and Francesca Hagethorn of the University of Amsterdam for their help and support. Thanks also to Anique Persoon who spent her Christmas holidays in the library doing a full set of experiments. We thank Soan Lan Ie and Peter Nugteren from the Bibliotheek Almere for their time, effort and enthusiasm they put into

this project. Finally we would like to thank the Waag Society⁴ for sharing their experience on game development. thank Alex de Vries for his C++ code that contributed to this project.

References

- [1] Libgaming google group.
<http://groups.google.com/group/LibGaming/>
- [2] Newgrounds.com, webcam games.
www.newgrounds.com/game/webcamgames
- [3] Opencv software library.
www.opencv.org
- [4] Vivid group inc.: Gesture xtreme system, 1986.
www.vividgroup.com
- [5] Reality fusion inc: Gamecam, 1999.
www.vividgroup.com
- [6] Setpixel.com: Magic coloring wall, 2002.
www.setpixel.com/content/?ID=twothingsdotone
- [7] Sony computer entertainment inc.: Sony eye toy, 2003.
www.eyetoy.com
- [8] Sega superstars, eyetoy games, 2004.
www.sega.com/gamesite/segasuperstars/
- [9] Gaming lan-party in the library of vlissingen, netherlands, 2006.
www.vlissingen.nl/web/show/id=291737/contentid=12881
- [10] Microsoft xbox live vision camera, 2006.
www.xbox.com/en-US/hardware/x/xboxlivevision/
- [11] Nintendo wii, 2006.
<http://wii.com>
- [12] B. Avery, W. Piekarski, J. Warren, and B. H. Thomas. Evaluation of user satisfaction and learnability for outdoor augmented reality gaming. In *AUIC '06: Proceedings of the 7th Australasian User interface conference*, pages 17–24, Darlinghurst, Australia, Australia, 2006. Australian Computer Society, Inc.
- [13] S. Beauchemin and J. Barron. The computation of optical flow. *ACM Computing Surveys*, 27(3):433–467, 1996.

⁴Website: <http://www.waag.org>

- [14] J. Chen. Flow in games (and everything else). *Communications of the ACM*, 50(4):31–34, 2007.
- [15] M. Csikszentmihalyi. *Flow: The Psychology of optimal experience*. Harper and Row, 1990.
- [16] H. Desurvire, M. Caplan, and J. A. Toth. Using heuristics to evaluate the playability of games. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1509–1512, New York, NY, USA, 2004. ACM Press.
- [17] J. Eisenstein and W. E. Mackay. Interacting with communication appliances: an evaluation of two computer vision-based selection techniques. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1111–1114, New York, NY, USA, 2006. ACM Press.
- [18] M. A. Federoff. Heuristics and usability guidelines for the creation and evaluation of fun in video games. Master's thesis, University Graduate School of Indiana University, 2002.
- [19] M. Heerink, B. Kröse, B. Wielinga, and V. Evers. Enjoyment, intention to use and actual use of a conversational robot by elderly people. In *Proceedings of the third ACM/IEEE International Conference on Human-Robot Interaction*, Amsterdam, 2008. ACM.
- [20] A. H. Jorgensen. Marrying HCI/Usability and Computer Games: A Preliminary Look. In *Proceedings NordCHI'04*, pages 393–396, Tampere, Finland, Oct. 2004. ACM.
- [21] M. W. Krueger, T. Gionfriddo, and K. Hinrichsen. Videoplacement: an artificial reality. In *CHI '85: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 35–40, New York, NY, USA, 1985. ACM Press.
- [22] S. Laakso and M. Laakso. Design of a body-driven multiplayer game system. *Comput. Entertain.*, 4(4):7, 2006.
- [23] J. Levine. Gaming & libraries: Intersection of services. *Library Technology Reports*, 42(5), 2006.
- [24] T. W. Malone. Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games. In *Proc. 1982 Conference on Human Factors in Computing Systems*, pages 63–68, New York, 1982. ACM.
- [25] S. Reeve. The interest-enjoyment distinction in intrinsic motivation. *Motivation and Emotion*, 13(2):83–103, 1989.
- [26] K. Squire and H. Jenkins. Harnessing the power of games in education. *Insight*, 3(5):2–30, 2003.
- [27] P. Sweetser and P. Wyeth. Gameflow: a model for evaluating player enjoyment in games. *Comput. Entertain.*, 3(3):3–3, 2005.
- [28] P. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2002.
- [29] S. Wang, X. Xiong, Y. Xu, C. Wang, W. Zhang, X. Dai, and D. Zhang. Face-tracking as an augmented input in video games: enhancing presence, role-playing and control. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1097–1106, New York, NY, USA, 2006. ACM Press.
- [30] M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting faces in images: a survey. *Pattern Analysis and Machine Intelligence, IEEE*, 24(1):34–58, January 2002.
- [31] Z. Zivkovic. Optical-flow-driven gadgets for gaming user interface. In *Entertainment Computing ICEC 2004*, pages 90–100, 2004.