ChildCrowds

A UNITY3D CROWD GAMING FRAMEWORK FOR CHILDREN

Tommy Paaske Hansen | Medialogy 10th Semester | August 30th, 2013

Foreword

The following report is a written documentation of a Master Thesis carried out at the Education of Medialogy – Aalborg University Copenhagen in the spring of 2013.

The report is divided into 8 main chapters; Introduction, Pre analysis, Problem statement, Analysis, Design, Implementation, Evaluation and Conclusion.
Each chapter is divided into a number of smaller sections and subsections that cover various topics, within each chapter theme. Each chapter also comes with a short introduction, describing its content and purpose.
Even though chapters may be read independently of each other for specific insight into their individual themes in relation to the project, it is recommended that the report be read in its entirety, in a front to back manner for the best understanding of the project and the reasoning behind it.

Throughout the report a number of appendixes, which are of relevance to the project, will be referenced. The appendixes can be found on the accompanying DVD, enclosed at the back of the report.

Literary references are given in the sixth edition APA style with the authors last name, followed by the year of publishing enclosed in parentheses, like so; (author, year).
A complete bibliography along with a list of all tables and figures, can be found at the back of the report.

I would like to thank the staff and children of *Sundby fritidshjem* for their help and participation in testing the framework. I would also like to thank *Alexander Sasha Popovich (AlexP)* of Codelaboratories.com, for providing drivers for the Playstation3 Eye camera hardware used in the project, as well as being helpful on their forum.

Contents

1 Introduction

"Though the mass audience can be used as a creative participating force, it is instead, merely given packages of passive entertainment" – Marshall McLuhan (McLuhan, 1967)

In this quote from his 1967 book "*The media is the message"*, visionary media Guru Marshall McLuhan describes well, both the nature of most mass audience entertainment products, while simultaneously acknowledging the power hidden in audience participation.
So where am I going with this, let us start with some background on were this project originates from.

In the year2004/2005 I attended an 8 month stay at *The European Film College (EFC)* in the windy hills of Ebeltoft Denmark. The EFC is a Danish folk high school, which means that apart from offering courses on film related subjects, a stay at the school comes complete with shared living quarters, daily meals in the dining hall, duties and chores. I had a great time at that school, within the comfort of school related projects, I got to touch lightly on the many roles involved in professional film and TV production, but most importantly I got to meet a lot of interesting and creative people. Here 8 years later, in 2013 many of these people have gone on to study or work within various areas of the film and TV business and I stand to write my master thesis in Medialogy with a specialization in games, at Aalborg University in Copenhagen.

The reason I have chosen to start my introduction on this rather nostalgic note, is that my relationships from my stay at the EFC, should come to play an important role in the choosing and development of my master thesis subject.
A few months before I started working on this thesis, I was invited by one of my previous co-students of the EFC, to attend a screening of a project he had been working on at the *National Film School of Denmark*, where he is currently studying as a producer.
He and a director-student from Sweden, had been working on a project they call *Cinema Dell Arte* (Madegård & Balslev, 2013). Cinema Dell Arte is a project featuring live animation theatre for children, that is; the project uses live motion capture theatre-actor performances, mapped onto 3D avatar characters, in a 3D generated environment, in real time. This provides an experience where the live performance part lets the child audience interact directly with the characters of the ongoing narrative, while the use of 3D generated avatars and environment, allows for cheap and easily implemented, creative fantasy inspired settings and special effects.
After the show, I talked to the people behind this project, who expressed their desire to expand on the concept, including an interest in so called "crowd gaming" of which they had heard, but knew very little.
In short, crowd games allow groups or crowds of people to collaboratively interact with the game, "as one". This is in contrast to e.g. traditional single or multiplayer games where individual players each interact independently with "their own" game, playing either against a computer opponent or against each other.

After a few meetings, discussing the possibilities, I decided to devote my master thesis to developing a framework for crowd gaming for children. The framework should serve as an interaction platform onto which crowd games can more easily be developed, making the concept of crowd games more accessible to projects like that of Cinema Dell Arte.
This report serves as a documentation of the analysis towards such a framework, as well as the design, implementation, test and evaluation of a prototype of both the framework itself, as well as a few games within it.

1.1 Initial problem statement

The initial introduction, forms the basis for the following initial problem statement, from which the preliminary analysis for this project, will take its form:

"Is it possible to create crowd gaming experiences, specifically designed for children?"

To elaborate on what is meant by a "crowd game", I have here included my own definition of the term:

"A Crowd Game is a game that has been specifically designed to, as well as technically implemented for accommodating input somehow derived from the collective efforts of a crowd of players, and based on this input generate audiovisual output."

This initial problem statement leads directly to the preliminary analysis, which will further investigate its main topics:

- Crowds and crowd behavior (from a sociological perspective)
- Children (with a specific focus on child play, and play activities)
- Games (a quick overview of games as a term, game content and genres)
- Crowd games (a State of the art literature review of existing projects regarding interactive audience participation in general)

2 Preliminary analysis

Following the themes introduced in the initial problem statement, this preliminary analysis chapter contains sections on crowds and group behavior. Next is a section on basic theories of children and play behavior. The chapter concludes with a section on games and game genres in general and the chapter concludes with a state of the art literature review of existing crowd game projects.
As a whole the themes investigated in this chapter, will form the basis for a Final Problem Statement, which will shape the rest of the project.

2.1 Crowd and group behavior

The term crowd or group may simply be regarded as relating to "*A large number of persons gathered together*", or *"A number of people attending the same public function; an audience"* (Farlex, 2013). However, from a sociological perspective, quite a lot of work has gone into understanding and defining the concepts of crowds and groups, as well as the psychology of crowds and groups as a whole, compared to that of the individuals within them.

2.1.1 What constitutes a group or a crowd?

There is no universal definition of what exactly makes a group, a group. Since groups have varying sizes, structural forms, ways of communicating and interacting with each other. Generally one may say that crowds are merely large groups, or groups merely small crowds. The terms will therefore be used interchangeably throughout this section.
Although no universal definition exists, a few basic requirements can however be identified as contributing to being able to recognize a number of individuals as being part of what might be regarded as a group (Schacter, Gilbert, & Wegner, 2009)

- *Interdependence*: For an individual of a group to succeed in their part of the task with which a group is faced, they depend (at least to some degree), on the behavior and actions of other group members.

- *Social interaction:* For a member to accomplish his or her goal within the group, some amount of communication (verbal or non-verbal) is required to take place between the members.

- *Group awareness*: All members of a group must know and accept that they are part of that group, for them to be actual members.

- *Common goals:* All members of a group share common goals, and work collectively towards them.

In addition to these points, groups or crowds can have different structures, they may have one or several leaders, as well as members that follow a more "go-with-the-flow" type of behavior, or a group may divide into smaller sub groups each handling part of the shared goal.

Generally research suggests that the definition of a group, highly rests on the notion of interdependence, and of the notion of people coming together to share and attempt to reach common goals.
Several scientists have developed different theories for trying to explain this crowd or group psychology, generally they all deal with the behavior and thought processes of individual crowd members and/or of the actions and tendencies of a crowd as a whole.

2.1.2 The origin, and classic theories of crowd psychology

The first notions of crowd psychology originated in the late 1800s, initially from studies of criminology and the activities and psychological nature of large gatherings of people, so called "mobs", during riots, strikes, demonstrations and events of similar nature, typically resulting from political disputes. Throughout the 1890s a series of theses, articles and essays were published, eventually inspiring the French physician *Gustave Le Bon* to write his book "*La Psychologie des Foules*" first published in 1895, which became a best seller (Bon, 2013).
Generally theories of crowd behavior can be divided into two different basic directions, *Contagion* and *convergence* theory.

2.1.2.1 Contagion theory

Contagion theory includes those of Le bon, and similar views, like that of *Sigmund Freud*.
In contagion theory, minds of the individuals within the crowd merges together, to form a collective unconscious, in turn increasing the enthusiasm of each individual within the group, while at the same time making them less aware of the actual true nature of their (possibly irrational) actions (Pick, 1995).
Although Le Bon's Theory, is one of the earliest explanations of crowd behavior, it is still widely accepted by many people, this is in spite the fact that a "collective mind" has not been documented by thorough systematic studies.

2.1.2.2 Convergence theory

Opponents of the collective unconscious theories argue that crowds does not take on a "life of its own" but is merely a product of the intentions and actions of its individual members (McPhail, 1991).
This concept is also known as *convergence theory*, and argues that the behavior of a crowd is not a product of the crowd itself, but is carried into the crowd by a convergence of like-minded individuals, that is people with common opinions, desires etc. come together to act out these in a crowd. Individuals of a crowd may sometimes do things that they would not have the willpower or courage to do on their own, since the very notion of a crowd shifts focus from the behavior of the individual, towards that of the crowd, which again can diffuse issues of responsibility.

2.2 Initial view on children and play

Many different theories exist within the field of children and play, be it from a social, psychological or educational perspective. In order to effectively design a crowd gaming experience suitable for children, some of these theories are initially investigated here. As with the State Of The Art review of crowd interaction systems, this initial view on (child) play, should provide enough background to again make educated decisions on the scope and direction of the rest of the project, following the Final Problem Statement.

2.2.1 Classic theories of play

We will begin by looking at some classic theories of what "play" is. We all know play when we see it, yet it is hard to precisely describe what constitutes play, through history many have tried such definitions, and as the fields of Psychology, Psychoanalysis and Cognitive psychology has grown larger through the 20th century, many of the older theories which were primarily grounded in philosophy rather than empirical research, have now been discarded or outdated.

2.2.1.1 The surplus energy theory

German philosopher *Frederich Schiller* (born 1759, died 1805), postulated the now outdated theory of play known as the "surplus energy theory", the theory suggests that:

"Play is an aimless expenditure of exuberant energy" – Frederich Schiller

The theory suggests that all living organisms generate energy, enough to cover basic survival needs, play in this optic is seen as a means of relieving the body of any energy that might be left over, after these basic survival needs have been covered.
The British philosopher Herbert Spencer, later adopted a separate version of this theory, he noted that at a very early age, the organism has a greater yearning for play than at an older age. Since younger members of a species typically depend on their parents for survival, they use less energy on these activities and therefore have more energy left over for play (Saracho & Spodek, 2003).
Generally throughout the history of psychology there have been a tendency to oversimplify complex behavior, the surplus energy theory of play is a good example of this (Schnell, 2008).

2.2.1.2 The practice or Pre-exercise theory

German philosopher and psychologist, *Karl Groos* (born 1861, died 1948) developed the theory known as *The practice or Pre-exercise theory,* it identifies a number of purposes of childrens play in relation to adult games, rituals and competitions. A key point in Groos theory, is that children through play are motivated to imitate and practice adult roles, and thereby play prepares children for adult life. (Saracho & Spodek, 2003)

2.2.1.3 Recapitulation theory

The theory known as *The Recapitulation theory* was Developed by American psychologist *G. Stanley Hall* (born 1844, died 1924), it is based on *Charles Darwins* theory of evolution and suggests that an individual reestablishes its species development. That is, each human originates at an embryo stage, and must go through evolutionary stages similar to those of the human species on their way to adulthood. Throughout this process, children dramatize the different stages (animal, savage, tribal member etc.) through play, and this play in turn relieves the child of any primitive instincts associated with each stage, that are not suitable for modern society (Saracho & Spodek, 2003).

2.2.2 Modern definitions of play

Although many of the older theories of play have today been discarded, they still form the basis, and provide a good point of reference and comparison for the modern theories. These modern theories are typically based around strong theoretical groundwork, supported by empirical research, and many of them place play as having an essential role in the psychological development of a child.

2.2.2.1 Psychoanalytic play theory

Grounded by Sigmund Freud and his scholars, psychoanalytic theory concerns dealing with, and being assisted to deal with suppressed conflicts that have occurred in previous life, and are now hidden in the subconscious.

In this perspective, Freud proposed that play performed a special and important role in the emotional development of children, play helps children deal with negative emotions and replace them with positive emotions. Freud also noted that children use play activities as a tool to better understand their own thoughts and emotions. (Saracho & Spodek, 2003)

Thus there is value in play in allowing children to express negative emotions that relate to situations which they have no control over in their everyday lives, these may include traumatic experiences and conflicts. (Verenikina, Harris, & Lysaght, 2003).

2.2.2.2 Arousal seeking/Modulation play theory

Professor of psychology *Daniel Berlyne*, proposed a theory of play suggesting play as a means of maintaining a balanced level of arousal, the function of play is here to raise or lower levels of stimulation, depending on whether a child is under- or over-stimulated. Play provides novelty, uncertainty and complexity in optimal doses, or levels, suitable for children. This balance between the new and the familiar is often seen applied to different methods within the field of education. (Berlyne, 1960)

2.2.2.3 Cognitive play theory

Swiss developmental psychologist and philosopher *Jean Piaget* shifted the focus of studies within the field of children and play, from social and emotional aspects, to childrens cognition. Piaget placed play within a previous theory of cognitive development and gave it a significant role in the development of the minds of children. In Piagets theory, play contributes to the cognitive development of children through two processes *assimilation* and *accommodation*. Assimilation being the dominant process in play, children tend to take something and make it fit to what they already know. This form of symbolic play, supports the importance of play in the development of childrens mental representations and abstract thinking (Piaget, 1962).

2.2.2.4 Sociocultural play theory

The Soviet psychologist *Lev Semyonovich Vygotsky*, places play as the most important leading activity in the early childhood years, that is the most significant psychological achievements and changes in a child occurs during activities of play.

Vygotsky also states that in situations such as make-believe play, children function above their own normal cognitive abilities, level of logical thinking, memory and attention. Their ability to display and self-regulate, deliberate behavior in make-believe play is also greater than in their normal everyday norm. Vygotsky also notes childrens ability to separate thought from actions and the division of objects into mental representations and symbolic functions.

Through pretend play, children are able to create an imaginative alternate dimension in which they can use substitution of objects and events. Separation of the meaning from the object supports the development of abstract thought and abstract thinking (Vygotsky, 1977).

2.3 Different levels and stages of play

When we discuss the concept of play, we also have to take into consideration the social context, age group and level of maturity of the child or children playing. The play of a toddler is significantly different from that of a preschooler or a teenager. Different theorists of play use different levels and stages of play to describe these differences, and to better understand how the concept of play changes along with the age of the child or children playing.

2.3.1 Mildred Partens six stages of play

American Sociologist *Mildred Parten*, was one of the first to conduct extensive research in the area of children and play, using anthropological field research methods, she studied children of various age groups in intense one minute periods and concluded that the play behavior of children could be categorized into 6 different stages – called *The six stages of play* (Parten, 1932)
The six stages of play presented by Parten are:

- *Unoccupied play:* Technically this stage does not involve play at all, and thus some controversy exist if there are actually five or six stages of play, Unoccupied play relates to children situated in the same area as other children playing, but not participating.

- *Solitary play:* Also known as independent play, relates to children playing by themselves, and maintaining this state by being focused on the play activity, ignoring or possibly not noticing the activities of other children around them.

- *Onlooker:* Similar to unoccupied play, the onlooker stage refers to children not actually engaged with play in a direct and active way. However, where the unoccupied child simply happens to be near other children playing, the onlooker may actually interact socially with other playing children by talking to them, asking questions related to the play activity without actually joining in.

- *Parallel play***:** Also known as adjacent play, relates to the instances where a child is engaged in the same or similar playing activity/activities as one or more children close to it. A child engaged in parallel play, will copy or mimic the activities of other children, but as with the onlooker, not actually engage in joined play.

- *Associative play:* This stage takes the focus a bit away from the playing activities and more towards the children engaged in them. Associative play, relates to when a child is actively interested in other children playing, but not the actual play activity itself, or when there is no actual organized activity at all. In this case the child merely wishes to be part of the playing group, a desire which again can be related to theories of self and group-identification.

- *Collaborative play:* This last stage covers instances where a child is interested in and engaged both socially with other children and actively with the playing activity. In

Collaborative play, the playing activity is organized, it is governed by rules (which may change during play), and participant have assigned roles (which may also change) (Santrock, 2007)

2.3.1.1 Mildred Partens six stages and their relation to age and maturity

Parten notes how the first four stages of play are typical of young socially immature children, the older the children get the more common activities further down the list become. The state of parallel play marks in some ways a threshold, where child begins to develop social awareness, leaning more and more towards a more social and collaborative playing behavior. Collaborative play is by far the most mature state of play, as it requires both a sense of self-identification as well as self-organization. Collaborative playing behavior can be expected from somewhere between the pre-school (2 – 5 years) and the beginning of the elementary-school–age (6-10 years), typical examples of collaborative play includes traditional "folk games" like *tag*, *tag the pole, freeze tag*, to name a few (Nieboer, 2013). The six stages are summed up in Table 1 on the following page.

While Partens theories are by now quite old, they are still widely accepted by modern scholars, however some controversy exists on whether the stages are truly "stages" in a chronological sense, and that children must traverse through one stage onto the next. Suggestions have been made about other factors influencing the playing behaviors and activities of children, including the level of maturity in the child, and how well the children know each other (Hughes, 2009).

Table 1: Sum up of Partens six stages of play in relation to social engagement, age and maturity

Stage of play	Typical playing activities	Age and social maturity level
Unoccupied play 	No play activity	No level of social maturity, the child is not engaged in play at all, but merely situated close to playing children.
Solitary play 	Playing alone with blocks or similar toys	No level of social maturity, the child plays alone ignoring others around it. Expected in the toddler to beginning of pre-school ages (½ - 3 years of age)
Onlooker play 	Observing other children play in various forms, actively asking into the concepts, rules and roles of the playing activity	Limited level of social maturity, the child does not partake, but may actively communicate with other playing children. Inquiring about concepts and rules. Expected in the pre-school ages (2-5 years of age)
Parallel play 	Playing with the same or similar toys and or actions as other children, but not actively playing along or communicating with these	Limited level of social maturity with no direct interaction or communication. Expected in the pre-school ages (2-5 years of age)
Associative play 	Many different playing activities, but with focus on being part of the group, rather than on play itself	Some level of social maturity, relating to both self and group-identification Expected at the end of pre-school (5 years of age)
Collaborative play 	Various playing activities containing advanced concepts of roles and rules. Activities can e.g. be pretend play, or folk games like *tag,* skip the rope etc.	Most socially mature state requiring both self-identification and self-organization Expected between end of pre-school (5 years of age) and beginning of elementary school (6 years of age)

2.3.2 Piagets three levels of play

Piaget described play in accordance with stages of cognitive development, thus play is by Piaget divided into three different levels:

- *Functional or practice play*: this level is associated with the sensorimotor stage and consists of repetitive motor movements with or without objects.
- *Symbolic/pretend/make-believe play*: When children are in the preoperational stage of development they start to engage in this level of play, *Vygotsky* further elaborated on this stage suggesting that there are two levels to symbolic play: play with objects or a simple act of pretend and symbolic role play.
- *Play/Games with rules:* This is when play is based on understanding and following rules, which is arguably related to the study of games, here play can occur individually or in a group (Piaget, 1962).

2.4 Games

This section will go into more detail of the concept of games, game definitions, genres and how these can be related to crowd games.

2.4.1 Game definitions

The actual game part of a crowd game have been defined by a large variety of traditional game scholars, the definitions vary, as does common personal understanding of what constitutes "a game". However many of the more accepted definitions have some things in common in this section several different definitions are analyzed and finally a list of their key points that constitute a game, is devised.

2.4.1.1 Elliot Avedon and Brian Sutton Smiths

"Games are an exercise of voluntary control systems, in which there is a contest between powers, confined by rules to produce a disequilibrial outcome" (Avedon & Sutton-Smith, 1979)

Elliot Avedon and *Brian Sutton Smiths* definition from 1979, is a powerful and very concise one, it is however also very scientific in its language, and would probably not find much use within the actual games industry today, however it contains some important statements.
First off, it describes games as something you engage in willingly, and something that you as a player has control over. Then it introduces the concepts of contest and power, meaning that as you play you compete somehow against someone or something in the attempt to reach some goal. This competition is governed by rules, and finally the game will result in a "disequilibrial outcome", meaning not only that one part wins and the other loses, but also that the competing powers can be viewed as having been "even" at the start of the game.

2.4.1.2 Tracy Fullerton, Chris Swain and Steven Hoffman

Tracy Fullerton, Chris Swain and *Steven Hoffman* have another definition:

"A game is a closed formal system that engages players in structured conflict and resolves its uncertainty in an unequal outcome" (Fullerton, 2008)

This definition has quite a few things in common with the first one. First off, Games are here described as a closed formal system, meaning that a game is made up of a number of interrelated components, that all have been clearly or "formally" defined. "Closed" means that the system of the game it is somehow limited by boundaries. This can again can be related both to the rules that govern the game, and what is and is not possible within the game system, but also to how you as a player "feel" when you are mentally "in the game" as opposed to when you are not. The definition introduces another new term, namely "engage", engagement and immersion is a whole different area of research, and many will argue that these are qualities of "good games", and not necessarily qualities that all games possess. From a point of view of game design however, it is an important goal to strive for. Finally this definition also stresses the point that the game is ultimately resolved in an unequal outcome. Once again there is a goal to reach, and a winner as well as a looser, be it the game itself or another player.

These definitions are created mainly from a scientific standpoint, in an attempt to fully enclose the term "game" in some "meaning". However when it comes to actually using them as guidelines for designing games, they will often be hard to put to practice,

2.4.1.3 Jesse Schnell

Several authors of game design books, have attempted to define games through lists of qualities derived from other definitions such as the ones above. *Jesse Schnell*, is one such author, and lists the following points:

- *Games are entered willfully*
- *Games have goals*
- *Games have conflict*
- *Games have rules*
- *Games can be won and lost*
- *Games are interactive*
- *Games have a challenge*
- *Games can create their own internal value*
- *Games engage players*
- *Games are closed formal systems* (Schnell, 2008)

2.4.1.4 Scott Rogers

Scott Rogers, another game design author, similarly sums up what a game is in list form, though much shorter.

A game is an activity that:

- *Requires at least one player*
- *Has rules*
- *Has a victory condition* (Rogers, 2010)

These lists provide a good basic understanding of what constitutes a game, of course creating a truly great game, requires both skill, experience and possibly even a bit of luck (at least from a commercial perspective).

2.4.2 Game genres and content

In his book Scott Rogers also include a very comprehensive list of different game genres, as well as the general game mechanics behind them (Rogers, 2010).
From the perspective of a crowd game, this may prove very important as different genres of games play very differently, and while some genres may be particularly good for crowd interaction, others may have game mechanics that directly or indirectly hinder it.

Scott Rogers list have been included here in a slightly edited form, the list may serve as a general overview of the different genres available in game design, and so that their affordances and constraints in relation to crowd interaction may later be reviewed.

2.4.2.1 Action games

The action game genre generally covers games which has hand/eye coordination as a key focus. The genre has many sub genres that each determine different content:

- *Action-Adventure*: featuring an emphasis on item collection and usage, puzzle solving and long-term story related goals. Examples include games like the *Tomb Raider* and *Prince of Persia* series.

- *Action-Arcade:* are games presented in the classic style of early arcade games, this genre has an emphasis on quick "twitch" gameplay, (high)scoring and relatively short playtimes. Examples include *Dig Dug, Donkey Kong, Diner Dash* etc.

- *Action-Platformer:* platform games typically feature a character running, jumping, swinging or bouncing his way through challenging "platform" environments, shooting and fighting enemies may also be involved. A crown platform-game example is Nintendos *Mario* series.

- *Stealth-Action:* is a sub-genre of action games that has a focus on stealth elements like hiding, sneaking as silently dispose of opponents. Examples include the *Metal Gear* and *Thief* series.

- *Fighting-Action:* an action game genre where two or more opponents battle each other in arena like environments, they are distinguished from other action games by the depth of their character controls, which typically include a wide variety of special moves and combinations. Examples include the *Street Fighter* and *Mortal Combat* series.

- *Beat 'em up/hack n' slash action:* these games have the player battle wave after wave of increasingly difficult enemies using simpler moves than those of Fighting-Action games.

2.4.2.2 Shooting games

Shooting games, or so called *Shooters*, focus (as the genre name reveals) primarily on firing projectiles. While typically being fast phased and "twitch" minded like action games, the genre has evolved into its own, including a number of sub genres that can be distinguished by their camera view:

- *First person shooters:* are shooters where they game is seen from the perspective of the in-game character, the tight camera is more limiting but also more "personal" than other shooters. Classic examples include the *Doom* and *Quake* series.

- *Shoot 'em up:* these games are arcade-style shooters, where players shoot large quantities of enemies while avoiding different hazards. The players avatar is typically a vehicle (like a spaceship), rather than an actual character. Shoot 'em ups can be presented from a number of different camera angles. A classic game example is *Space Invaders.*

- *Third person shooters:* a shooter where the camera is locked to a position further behind the game character, giving a larger view of the character and his/her immediate surroundings. Despite the wider view, emphasis remains on shooting. Examples include the *Star Wars Battlefront* and *Max Payne* series.

2.4.2.3 Adventure games

Adventure games focus on puzzle solving and item collection as well as inventory management and sometimes character optimization. While the first adventure games in history were entirely text based, this overall genre now includes a variety of different sub-genres:

- *Graphical Adventure (Pont and Click Adventure):* in this sub-genre, players use their mouse to navigate the game world by clicking where to go, the mouse is also used to find items, uncover secrets and other ways of interacting with the game. Examples include the *Leisure Suit Larry* and *Monkey Island* series.

- *Role-playing Adventure:* This game genre is based on classic pen and paper role playing games like *Dungeons and Dragons*, it involves choosing one or several characters, a character classes, skills, abilities, choice of weapon proficiencies etc. adding to and increasing these throughout the game, through combat, exploration, treasure finding etc. Examples include The *Baldurs Gate* and *Eldar Scrolls* series.

- *Survival/horror Adventure Games:* focus on surviving in a horror scenario with sparse equipment, weapons ammunition etc. Examples include The *Resident Evil* series and *Amnesia: The dark Decent.*

2.4.2.4 Construction and management games

In this game genre players build and manage complex concepts like a zoo, theme park or entire cities. Players have limited resources and typically play with the goal to constantly optimize and make their creation larger, while following smaller stretch goals. Examples include the *Sim City, Theme Park* and *Tycoon* series.

2.4.2.5 Life simulation games

Similar to construction and management games, life simulation games involve building and managing virtual relationships. Examples like *The Sims series* combines the two by letting players both build houses for their *Sims* while at the same time managing their relations to neighbors, friends, partners etc.

- *Pet simulation games*: are a sub-genre of life simulation, and originates in the 1990s Japanese *Tamagotchi* trend. Instead of nurturing human characters, the player nurtures a virtual pet. Examples include the game *World of zoo.*

2.4.2.6 Music and rhythm games

This game genre uses sound as a main game mechanic, in that it typically has players trying to match rhythms, beats or musical scores in an attempt to reach high scores. The games differ from other genres as they often include specially designed physical interfaces like dance mats, microphones or controllers in the shape of instruments. *Singstar, Dance Dance Revolution* and *Guitar Hero* are good examples of this.

2.4.2.7 Party games

While maybe a bad genre title, the party game genre covers games that are specifically designed for parties and similar social events. Gameplay is often competitive or team based and typically in the shape of various mini-games with very short playtimes. *Mario Party* and *Buzz!* are examples of party games.

2.4.2.8 Puzzle games:

While other game genres like some Action or Adventure games sometimes contain puzzle elements, the puzzle genre covers games whose main focus and mechanic revolves around puzzle solving. The can be slow and methodical or fast paced requiring quick thinking and decision making. Examples of slow puzzle games are *The Incredible Machine* and various *block-moving* titles, while fast paced puzzle games include *Tetris* and *Pipe dream* as good examples.

2.4.2.9 Sports games

Technically sport is more of a theme than a genre, and sports games can have very varying content and mechanics. Generally they can be divided into two sub categories:

- *Sports simulation games:* in which the actual sport is simulated through gameplay, letting the player experience many of the same elements that exist in the actual sport. Examples include the *NFL: Madden* and *Tony Hawk* series.

- *Sports management:* as much a management as a sports game, sports management games focuses on managing the team of any team sport, choosing players, strategies, tournaments etc. Examples include the *Championship Manager* and *NFL Head Coach* series.

2.4.2.10 Strategy games:

While strategy games might still include an element of battle or combat, they have their focus shifted from fast twitch interaction to thinking and planning. Differences in how and when players interact with the games are what primarily distinguishes the different sub genres of strategy games:

- *Real Time Strategy (RTS) games:* have players issue commands to different units in real time, often the player will assume the role of a commander or general, building armies and issuing orders to different troops in battle. While having a focus on strategic planning and thinking, these games can at times be quite fast paced. Examples include the *Command and Conquer, Age of Empires* and *StarCraft series*.

- *Turn-based Strategy games:* these games have a slower phase than RTS games, allowing players to really plan a strategy and think about moves before, in turn, submitting them. The genre draws inspiration from and includes adaptions of many non-digital board games like for instance Chess. Digital examples that are not based on board games include the *X-Com* series and *Frozen Synapse.*

- *Tower defense:* is a special case strategy game that resembles themes from construction and management games as well. Here players create automated projectile shooting structures to keep increasingly difficult waves of enemies at bay. Examples include *Defense Grid* and *Plants vs. Zombies.*

2.4.2.11 Vehicle simulation games:

Like it was the case with some sports games, vehicle simulation games are designed to simulate a real experience, in this case piloting a vehicle. Emphasis is here placed on making the experience as “real” or realistic as possible, even when handling spaceships that does not exist in reality, the feeling should still prove somewhat realistic. There are two basic sub-genres of vehicle simulation, each relating to a separate type of vehicle:

- *Driving simulation*: players “drive” different grounded vehicles from motorcycles and speed boats to race cars or monster trucks, around on tracks. Since many of these vehicles exist in reality, these experiences can be made ultra-realistic. Examples include the *Need For Speed* and *Motorstorm* series.

- *Flying*: apart from adding another dimension of control (up and down) flying simulation are in many ways similar to driving simulation. Many flying simulation games however include experiences that players are less likely to have any real life relation to, like for instance that of flying a spacecraft or super jet fighter. Examples of flying simulation games include the *X-Wing* and *Flight Simulator* series.

While it is hard to cover all possible games in terms of game genres, and many games today are a mix of many game genres throughout the gameplay, by far the majority of games that exist today can however be fitted into one or more genres on the list presented here.
What is immediately apparent is how different the genres are in game playing mechanics and content, and in turn how different the actions required of the player, are.

2.5 State Of The Art crowd gaming and interactive audience participation

This section will explain the term "crowd game" in depth, by analyzing the different methods of approach for input and output through existing systems designed to be interacted with by a crowd. The section will present and a number of relevant State of the art Crowd Interaction projects, designed to be interacted with by a crowd of people.
The analysis is done primarily with a focus on the input methods chosen, with the intent to later make an educated decision on the choice of approach for this particular project. The section will also briefly explain the venue and game or other experience associated with each presented case, for contextual reference as well as mention any findings and/or advice presented in the project documentations

2.5.1 Cinematrix

While quite old by now, we will begin by looking at the crowd gaming system presented by *Loren* and *Rachael Carpenter* as pre-show entertainment at SIGGRAPH in 1991; *The Cinematrix system* (USA Patentnr. #5210604, #5365266, 1991). Since its reveal it has been a source of constant reference throughout other documented projects that have been developed, and Cinematrix is still in use today under ownership of Cinematrix Inc. (Cinematrix, 2013)

2.5.1.1 The Cinematrix system

The Cinematrix system worked by issuing a large number of paddles to the audience. Each paddle had a red and a green light reflective side that could be picked up by a camera situated by a large screen and facing the audience. Depending on whether the majority of the crowd showed the green or the red face of the paddle, the associated media experience reacts in a specific way.
No actual documentation of the specific techniques behind the Cinematrix system exist, however judging from descriptions and pictures of the setup, the green and red colors on the paddles are of a particularly reflective nature, and responds well to bright light being shone upon them (See Figure 1).

Figure 1: An audience interacting with the Cinematrix system – From the Cinematrix website (Cinematrix, 2013)

From here, segmenting the red and green paddles from a continuous video stream is a simple matter of color thresholding possibly followed by some noise reduction.

These subjects are investigated further in a later chapter of this report (See Section 4.1 - Computer vision techniques).

2.5.1.2 Cinematrix venue and use

The SIGGRAPH 1991 demonstration of the Cinematrix system took place in The SIGGRAPH Electronic Theatre, and was used for several different things:

- Opinion polling: In which binary choices presented on a large collective screen, are decided by, based on the majority color presented by the audience.
- Maze navigation: in which the audience collaboratively have to steer left and right through a maze, by showing either the green or red side of the paddle.
- Racing games, in which the audience steers a race car left or right by showing the green or red side of the paddle

Several other games are displayed on the Cinematrix website, including, amongst others a "*light sabre fight",* a "*Flying game*" and a *"Basketball game"* all of which use similar input methods as described above (Cinematrix, 2013).

2.5.2 Squidball: An experiment in Large-Scale Motion Capture and Game Design

Following the success of the Cinematrix system, a project for SIGGRAPH 2004 was developed, using the same venue (the Los Angeles Convention Centre) as pre-show entertainment for the SIGGRAPG Electronic Theatre, but using a very different technical approach.
Squidball (Bregler, et al., 2005), used a large scale motion captures setup as input for various crowd games that were played by up to 4000 people.

2.5.2.1 The Squidball system

The Squidball project used a Vicon motion capture system (VICON, 2013) with 22 MCAM2 cameras, each with a field of view of 60 degrees and a 1280x1024 pixel resolution.
By physically scaling up the reflective ball-markers traditionally used in motion capture, by a factor of 100, the team behind the project were able to capture the balls at a distance of 250 feet (approximately 76 meters), compared to the normal 25 feet capture range using normal sized ball-markers.
For the scaled up ball-markers the team used helium filled weather balloons fully covered with M3 retro reflective tape.
They then used strong halogen stage lights to shine light on the balls and passed 12 balls to the audience, who could then bounce them around over their heads, as a method for interacting with a game running on a large collective screen on the stage in the front of the room (See Figure 2).

Figure 2: Electronic theatre audience playing Squidball (Bregler, et al., 2005)

2.5.2.2 Squidball venue and use

As mentioned the squidball system was introduced in the Los Angeles Convention center, in Hall K, with a total space of 240 x 240 feet (approx. 73 x 73 meters). A game, specifically designed for the system, was used.
The 12 reflective balls in physical space, were represented in the digital game space as green spheres on the screen. Players moved the balls around in Hall K, and their position corresponded to a position in 3D game space. The object was to hit and thereby destroy, changing grids of yellow spheres in 3 levels of increasing complexity.
Although generally regarded as a success, the documentation of the Squidball project notes a few challenges in and problems in their design:

- The system is hard to set up and requires a carefully planned calibration session which for the 2004 SIGGRAPH presentation took about 30 minutes (for the calibration alone)

- Players of the game often had a hard time concentrating on looking at the screen (which is in one direction) and at the same time bouncing a ball in another direction to make it hit its targets

- The system is generally expensive, both in terms of Motion capture equipment and the specially designed helium filled weather balloons.

2.5.3 Techniques for Interactive Audience Participation

In the 2002 paper "*Techniques for Interactive Audience Participation*" (Maynes-Aminzade, Pausch, & Seitz, 2002), 3 techniques for having a large audience interact collaboratively with content on a large screen, was explored. The techniques were tested with several different types of games in 30 demonstration sessions with participants of between 150 and 600 movie audience college students, all of the techniques makes use of a single camera, and are intended for a seated audience (like that of a cinema).
Each technique and the corresponding game activity will be described here, followed by a sum up of the team findings.

2.5.3.1 Audience movement tracking

This technique builds on the actual movement of the audience themselves, without any additional components, using a template matching method. The camera faces the audience, and at the beginning of a session, 3 reference images are stored of the audience. One where the audience sits normally in their seats, and one with them leaning left and right respectively. For each frame of the continuous video feed provided by the camera during a play session, the frame is compared to each of the 3 reference images, using a sum of squared differences approach. The resulting correlation coefficients are used to compute the continuous control value used as a parameter for the onscreen game.

For the audience movement tracking, the classic game of pong (See Figure 3), as well as games within the racing genre was used. These corresponded well with the human natural tendency to move left and right to express a desire to move an object or steer a car.

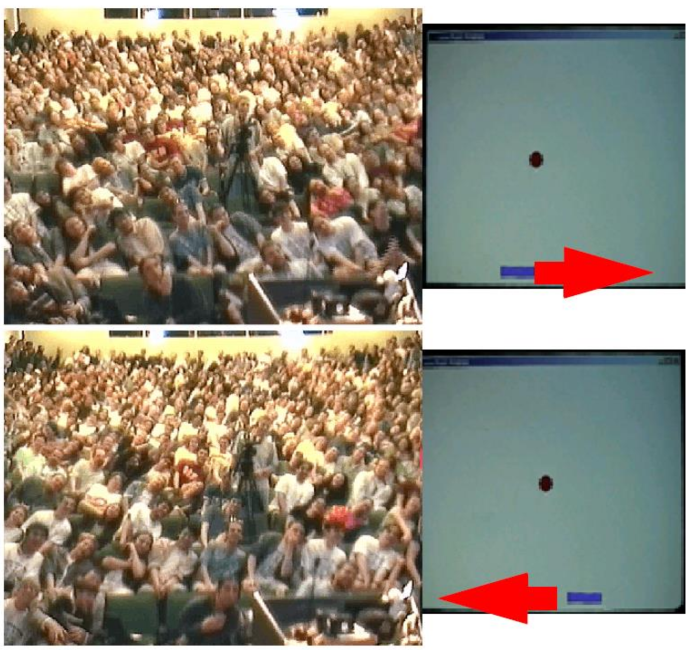


Figure 3: A large audience leaning left and right to control the paddle in a game of pong (Maynes-Aminzade, Pausch, & Seitz, 2002)

2.5.3.2 Beach ball shadows

The second technique presented in the paper is based on tracking the shadow of a beach ball onto the large game screen. The camera here points to the screen, and computer vision techniques such as thresholding and noise reduction is used to segment out the largest round black dot on the screen. The position of this dot is used as a “mouse pointer” in the ongoing game.

The beach ball shadow technique is documented as having been tested with the classic game of missile command, in which the player(s) must stop descending missiles from destroying the landscape at the bottom of the screen. This is done by batting and bouncing the beach ball around, making its shadow hit the missiles (See Figure 4).

Figure 4: The shadow of a beach ball acting as a cursor in a game of Missile command (Maynes-Aminzade, Pausch, & Seitz, 2002)

2.5.3.3 Laser Pointer Tracking

The final technique presented in the paper is based on tracking of laser pointers, pointing at the screen. Here the camera also faces the screen, and again Computer Vision techniques segments out the bright dots formed by the laser pointers on the large screen, these can then be used for a number of different purposes depending on the game or activity at hand, the paper documents this method as having been tested with a number of different applications, these include:

A collaborative painting program; in which each dot has a line drawn behind it allowing the audience to paint on the screen using different colors for each laser pointer (See Figure 5).

*A version of the game "Whack-a-mole***;** in which the audience must rapidly flock their laser points to hit small moles that appear randomly on the screen (See Figure 6).

A "Scratch and reveal"; in which the audience must use their laser pointers to gradually remove a black layer covering some humorous image hidden beneath (See Figure 7).

Trivia competitions; in which the audience "votes" for what they think is the correct answer to a question, by shining their laser dot on in on the large screen. The answers that has the most votes when a predetermined time limit runs out, is chosen (See Figure 8).

Figure 5: The collaborative painting program, here the laser pointers appear as small red dots in front of each colored line

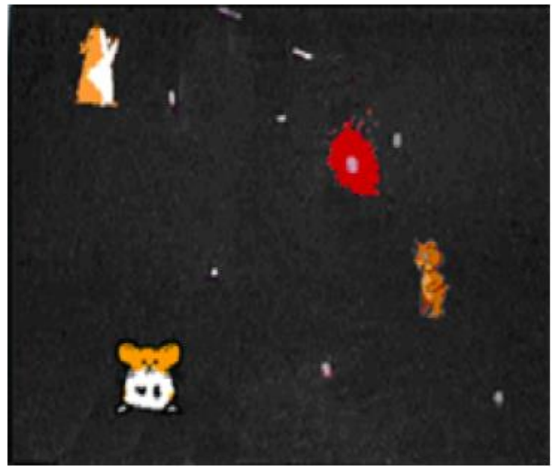


Figure 6: The whack-a-mole game with laser points scattered around the screen (small dots) amongst a number of moles

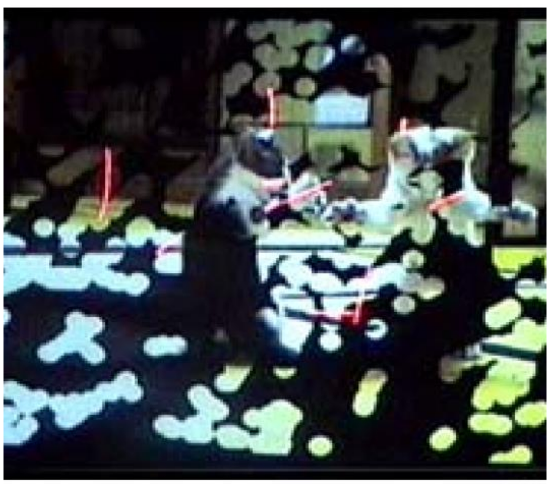


Figure 7: The "Scratch and reveal" game with the laser pointers appearing clearly as red stripes on the screen

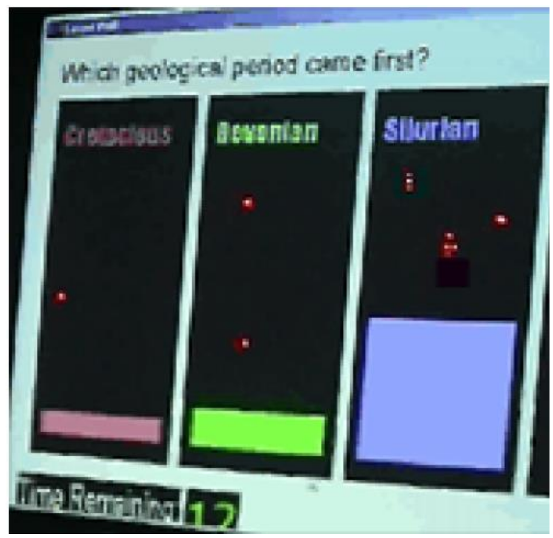


Figure 8: The trivia competition with a clear favoring of the third answering option. The time remaining is visible in the lower right.

Other games included a version of *"connect the dots"*, and *"form a figure".*

2.5.4 Engaging the Crowd – Studies of Audience-Performer Interaction

Engaging the Crowd – Studies of Audience-Performer Interaction (Barkhuus & Jørgensen, 2008), presents "cheering meter" technique for allowing the audience at a free-style rap battle concerts/competitions to "grade" the performers, allowing a winner to be declared based on the collective cheering efforts of the audience. No crowd game is involved in this system, but the input method is interesting, yet simple, and might also be used for crowd gaming purposes.

2.5.4.1 Cheering meter system

The system consists of 4 microphones recording at 44100 Hz, 16 bit and 1 channel. A number of 1/128 second samples would be taken, and measuring the volume of the samples each higher sample value would override any previous ones. Resulting in a "peak volume" measure. The measure was continuously translated to a score from 1-100 and displayed on 4 large LED displays surrounding the scene.

2.5.4.2 Cheering meter venue and use

The venues used for testing were a touring festival with the scene (a boxing ring) and equipment needed to be set up each day. The venue could be wholly surrounded by audience on all sides, which gave the reason for the 4 microphones and 4 LED displays (one for each side of the ring). After a performance from a rapper, a presenter would ask the audience to "*make some noise"* and the end measure on the LED displays would determine the outcome of the competition.

2.5.5 WeINteract – A Pervasive Audience Participation System

This paper from the 2004 CHI student competition introduces the "WIN" system – a method for capturing and quantifying an audience's natural performance in approval activities (such and waving and clapping) and using this measure to give an indication for an audiences grading of athletes during the Olympics (Chandrasekaran, Mohan, Pathipaka, & Saxena, 2004)

2.5.5.1 WIN system

The paper investigates several different techniques for implementation, but settles on using so called "wearables", small sensor fitted objects that can be worn as necklaces, wrist-bands, etc. or fitted into clothes.
In the case of the presented project, the wearable is a wrist-watch embedded with sensors for measuring motion and vibration, which translates into waving and clapping. The watch sends this data wirelessly to a central server that keeps track of performers and performances and translates the data to a score value that is sent back to the watch. The user is presented with the value and given an option to change it, before finally submitting.

2.5.5.2 WIN venue and use

The WIN system, is described as being intended for Olympic performances, and apart from the grading of athletes functionality, also presents two games for intended use with the system:

- *A training game***:** in which audiences are taught how to use the system, by having to grade the individual performances of two mascots on large collective screens.

- *A chariot racing game:* a game designed to be played prior to or in between performances, the stadium is divided into sections, and the collective insensitivity of clapping and waving of each audience section, determines the speed of their chariot.

2.5.6 BannerBattle – Introducing Crowd Experience to Interaction Design

BannerBattle, was a project proposed by two students of Aarhus University in Denmark, as a means of “gamifying” existing behavior of crowds at national football matches. (Veerasawmy & Iversen, 2012)

2.5.6.1 BannerBattle system

BannerBattle consists of two eight-meter long digital advertising-banner displays. One Banner facing the home-fans and the other facing the away-fans. Each crowd of fans has their audio and video recorded by a camera and a directional microphone. Each video feed is overlaid with the recorded teams colors, and the recorded audio output from the microphones was analyzed and visualized as an equalizer.

2.5.6.2 BannerBattle venue and use

Intended for a stadium, with an ongoing football match, each match would start with the banner having half its area covered by a different video feed, the two feeds separated by the equalizer visualization.
As a crowd cheers and moves physically, the intensity of these activities are measured in the video and audio recordings, and the crowd with the highest values will start conquering a larger part of the banner space (See: Figure 9)

Figure 9: One of the digital advertising banners of BannerBattle showing the team with blue colors cheering the most (Veerasawmy & Iversen, 2012)

2.5.7 MobiLenin – Combining A Multi-Track Music Video, Personal Mobile Phones and A Public Display into Multi-User Interactive Entertainment

MobiLenin (Scheible & Ojala, 2005), is a crowd interaction system for interacting with a multitrack music video in a case of joint authorship, using ordinary personal mobile phones. The system was developed with an intention of promoting more social interaction between people.

2.5.7.1 MobiLenin system

Like many other crowd interaction systems the MobiLenin setup consists of 3 sides, a client a server and a collective display. A server runs a music video on a large screen, and sends queries via HTTP and GPRS for crowd participants to vote on a desired next cause of action they would like to see from the music video actor. Six different synchronized tracks exists, and the track that receives the most votes is played next.

2.5.7.2 MobiLenin venue and use

MobiLenin was designed for use in cafes, restaurants, bars and similar venues where a music video may run in the background. Apart from providing the audience with a sense of joint authorship over the music video, as an incitement to participate, the system used a lottery system to draw out random winners of small prices (like beer or pizza) to participants.

2.5.8 Space Bugz! – A Smartphone-Controlled Crowd Game

Technically one could argue that despite the title of the paper in which it is presented, Space Bugz (Birke, Schoenau-Fog, & Reng, 2012) is not an actual crowd game, but merely a multiplayer game played on a large collective screen. However since it, in so many areas, resembles the setup of crowd games as otherwise investigated, it has been included here to consider the possibilities of using smartphones as input devices.

2.5.8.1 Space Bugz! System

Space Bugz! Is a classic arcade game in the style of Asteroids (Atari, 1979) , in which a player steers a spaceship from a top down view, avoiding obstacles and shooting enemies. The system consists of a game server, running the game on a large collective screen, players can utilize a specifically designed smartphone application that connects to the server and lets the player steer his or her individual “space bug” by tilting the smartphone. Players can shoot in different directions using a swiping finger motion on the smartphone display. Players can also produce a “pulse” that lights up their avatar on the large screen, making it easier to locate.

2.5.8.2 Space Bugz venue and use

Space Bugz! Was tested in cinema settings with up to 28 players playing simultaneously in the same game. Since the game does not support any collaborate interaction, players play to defeat each other in terms of points gained by shooting opponents in a death-match scenario.

2.5.9 Large Audience Participation, Technology and Orchestral Performance

Glimmer (Freeman, Large Audience Participation, Technology and Orchestral Performance, 2008) is a system first introduced in 2005, designed to engage a concert audience as musical collaborators by letting them interact with live orchestra musicians using battery operated light sticks.

2.5.9.1 Glimmer system

Glimmer consists of several elements; 4 consumer grade video cameras capture the audience turning on and off battery operated light sticks. A computer analyzes the video data and sends this to a separate computer that translates it into animation feedback as well as instructions to the live performing musicians. The musicians receives these instructions via multi colored lights mounted on their individual music stands.

2.5.9.2 Glimmer venue and use

Intended for live classical music venues, Glimmer has the audience divided into 7 different groups that controls a corresponding group of three of four musicians. Basically the percentage of the group that has their light sticks turned on, will decide the dynamic with which their corresponding group of musicians is currently playing. If everyone has their lights turned on at the same time, the group will play as loudly as possible, if everyone has it turned off, the group will be silent. At a higher level, the frequency with which an audience group changes the percentage of on and off light sticks, influence other characteristics of their musical group, like how often they play, their dynamic and the pitch at which they play. A large screen displays simple animations which helps the audience more easily understand the relationship between their actions and the music. Each group is represented by a rectangle, that changes color based on the group activities (see Figure 10).

Figure 10: An audience and orchestral collaboration using Glimmer (Freeman, Glimmer: Lights, Orchestral Performance, 2013)

Several issues were reported in connection with the project, primarily related to the short time for rehearsals between audience and musicians, and the inability of the audience within a group, to work together. This resulted in people within the group turning their light sticks on and off at random, effectively cancelling each other out. Also several audience members attempted to influence the music by waving and moving their light sticks, to no effect.

2.5.10 Composition for Conductor and Audience: New uses for Mobile Devices in the Concert Hall

Composition for conductor and audience (Roberts & Höllerer, Composition for Conductor and Audience: New Uses for Mobile Devices in the Concert Hall, 2011) is another crowd interaction system within the field of music and music composition.

2.5.10.1 Composition for Conductor and Audience system

The system consists of an iOS/Android application that the audience of a music concert can use on their smartphones. It utilizes an application software called control (Roberts, Control: Software For End-User Interface Programming, 2011), which is a general purpose solution for realizing touchscreen interfaces. *Control* allows for the app to communicate with a server handling the logic of the music composition as well as a visualization part of the system using Open Sound Control (OSC) (Wright, 2005). At the same time, the server is capable of sending instructions back to the app, queuing for the app interface to be changed dynamically throughout a music performance. The conductor of the piece holds another smartphone, also running a version of the Control application, the application tracks the conductors gestures utilizing the smartphones accelerometer.

2.5.10.2 Composition for Conductor and Audience venue and use

The system was designed to be used in connection with a custom concert performance, specifically composed to facilitate the corporation between audience and the conductor. After a brief introduction on how to participate, the conductor would cue the beginning of the piece and start directing the audience by slowly moving one arm up and down holding a smartphone, the audience members are instructed to move a simple virtual slider in response to these gestures. Up to eight members of the audience can participate, depending on the number of sound parameters exposed for control at any given time.

2.5.11 Giveaway Wireless Sensors for Large-Group Interaction

The final project included in this State Of The Art review, presents a low cost handheld wireless motion sensor for large group interaction (Feldmeier & Paradiso, 2004). The device is cheap enough to be given away at events for where large group interaction in the form of motion tracking is desired.

2.5.11.1 Giveaway Wireless Sensors system

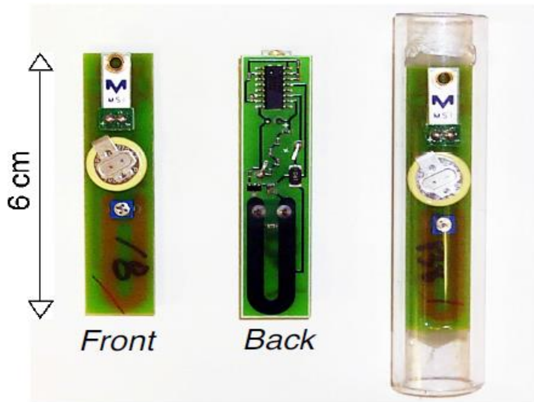


Figure 11: A cheap wireless motion sensor (Feldmeier & Paradiso, 2004)

The devices themselves consists of a piezoelectric foil accelerometer, a CMOS timer, and a single-transistor 300 MHz RF transmitter as well as an onboard battery (see Figure 11). The devices send a very short (50 µs) long, RF pulse when it is jerked, with a range of approximately 10 meters, making collisions of signals very unlikely, even when a larger audience tries to synchronize their movements. Signal can be derived into a set of features in real time, including activity levels (the number of hits arriving across different time intervals), significant events (more than a certain number of hits arriving simultaneously), and average tempo derived through Fast Fourier Transform of cross-correlation analysis.

2.5.11.2 Giveaway Wireless Sensors venue and use

Though not restricted to any particular activity, the system was tested on several occasions during “interactive raves” held at MIT with up to 200 distributed sensors. Rave dancers would hold a unit in each hand, and the derived features were mapped onto generated music, based on a set of rules. As such the music becomes more complex with increasing activity, the tempo produced is set to the detected tempo plus 2 beats-per minute, significant events with many closely-timed pulses produce a corresponding audio effect.

3 Pre analysis sum-up and Final Problem Statement

This chapter will present a short sum-up of the findings of the pre-analysis to provide an immediate overview before presenting the Final Problem Statement which will guide the rest of the project.

3.1 Sum-up of the pre analysis in relation to crowd games

In this section, the various topics investigated in the pre analysis are summed up, and related to the topic of crowd games. The topics are evaluated in relation to their relevance in designing a crowd gaming framework that can support playing activities for children, less relevant topics, methods and techniques are discarded, while the most important factors are picked out and condensed, for use in shaping the Final Problem Statement.

3.1.1 Sum up of crowds and crowd behavior

For a crowd game to actually become a crowd game, and not merely a number of people playing individually in parallel, the game must foster some or all of the requirements that make up a group or a crowd as identified in the pre analysis (see section 2.1, page 9).
This means that a crowd game must:

- *Support Interdependence*: Individual players depend on other players of the game to succeed.
- *Support Social interaction:* Individual players should be allowed or even encouraged to communicate (in a verbal or non-verbal manner) with each other in order to reach the games goals.
- *Create Group awareness*: All players should share a feeling of being part of the same group of players.
- *Create Common goals:* All players should share one or more of the same goals in the game, and be allowed to work collectively towards these.
- *Allow for different roles:* Individual player should be allowed to take on both a leader type role as well as "go-with-the-flow" type of role, depending on their preference.

The area of crowd psychology seems less important in the context of crowd games, as it typically seems to deal more with uncontrolled (and undesired) crowds of people, like those of a mob, riot, demonstration or protest. And what might motivate individuals within such crowds to perform illegal or irrational actions they would otherwise not have.
It should be noted however that through the design of a crowd game, focus can be shifted from

the individual player to that of a group, which might prove beneficial, especially when dealing with a group or crowd that possibly contains children of a shy personality.

3.1.2 Sum up of game genres and content

As the State Of The Art review shows (see section 2.5, page 24), when it comes to the genre or game content of crowd games, these may be specifically designed to accommodate the context in which it is played (like being related to a sports event, musical concert etc.), or simply be an adaptation of a simple well known game or game genre (like Pong (Atari Incorporated, 1972) or a Car Racing game). As such, crowd games does not require any specific rules or implementations different from that of several genres of traditional single player games. What is required however, is a source of input different from that of a keyboard or mouse, the input has somehow be a representation of the collective efforts of the crowd, for which several different methods exist. This also means that the actual gameplay has to be of a nature in which it can be understood and interacted with via the collective efforts of the crowd playing. This again means that the game cannot be of a genre in which personal strategy or playing style has too much of a central role in the gameplay.

Different game genres and content was briefly investigated in the pre analysis (see section 2.4.2, page 19), and revealed that different genres have hugely different game mechanics and content, as well as required playing styles. Relating these genres to crowd gaming, it becomes obvious that certain genres are more appropriate than others. Generally game genres may be divided into 3 categories in relation to crowd gaming:

3.1.2.1 Genres that work well:

Music and rhythm games: this genre is possible one of the best suited for crowd gaming, it has a focus on following beats, scores or tracks, and can for instance involve dancing along with a rhythm. An action that can easily be (and often is) partaken by an entire group of players.

Vehicle simulation games*:* these games typically involve obvious responses to events, like steering a car left or right to follow a race track. As such, a vehicle simulation game could easily be designed to accommodate the interaction from a crowd rather than an individual player.

Party games: are in their very nature great crowd games as they foster social interaction, competitive of team playing as well as short simple gameplay.

3.1.2.2 Genres that may work (depending on content):

Action games*:* depending on which sub-genre of action game, these might work as crowd games.

- *Action-adventure* with its focus on item collection and usage in puzzle themed challenges would work poorly, as no guarantee can be made that all players will solve a puzzle the same way or reach a solution at the same time, Also this sub-genres focus on long term story goals generates issues with finding a game pacing that suits all players.

- *Action-arcade* on the other hand could work well, with its main focus on quick twitch gameplay, with typically a single cause of action being the most appropriate.
- *Action-platformers* like *Mario* may prove to require too precise controls, and also provide a certain degree of freedom on where to go, which would have to be collaboratively decided on by the crowd.
- *Stealth action* similarly typically provides many possible approaches to the challenges and enemies within the game, and are therefore also best suited for the individual player.
- *Fighting action* has a main focus on deep controls and advanced possibilities of move combinations, which would be very difficult for a crowd to control.
- *Beat em' up* with a more simple control scheme than fighting games, could work by giving the crowd of players a single, or very few possible moves.

Shooting games: depending on the sub-genre a shooting game may work as a crowd game.

- *First person shooters* with their point Of View camera angles, are specifically designed to provide a more personal relation to the player, they also typically provide a wide number of options on where to move, take cover, engage enemies etc. and as such are a poor genre choice for a crowd game
- *Shoot 'em ups* like *Space Invaders* on the other hand could work well, due to their simpler controls, and limited possible causes of action.
- *Third person shooters* have many of the same disadvantages as First Person Shooters in relation to crowd gaming, even though the camera here removes some of the personal relation to the player avatar, they are similarly a poor choice of genre for a crowd game.

Sports games: Depending on the sport and the content of the game, sports games may work as crowd games, as with many of the other genres it comes down to the complexity of the sport and its simulation as well as the different possibilities of actions provided to the player(s).

- *Sports simulation*: can essentially have many of the same qualities as vehicle simulation games or similar genres, and for instance let the players guide a skateboarder or kite flyer through checkpoints. This type of gameplay could work well as a crowd game, while a simulation of a football or ice-hockey game do less so.

- *Sports management*: is closely related to the building and management game genre, and as such requires individual player strategies which corresponds badly with crowd gaming.

3.1.2.3 Genres that work poorly:

Strategy games: as they are mainly about beating the opponent using superior well planned out strategies, devised by individual players. Having to decide and agree upon the best cause of action for every move or action required in a strategy game, would most likely cause chaos.

Adventure games*:* with their emphasis on item collection, inventory management and storytelling, works poorly as they cannot be suited to the preferred rhythm and pacing of gameplay preferable to all players at the same time. This could eventually cause players to lose interest in the game, either because of boredom or confusion.

Management games: typically has a focus on the individuals personal strategy for optimizing and expanding on their creations, and as such lean towards the strategy genre

Life simulation games: are similar to the management games, which makes them a poor choice for crowd games for many of the same reasons.

Puzzle games: similarly to strategy games, require a lot of thinking and typically individual methodic approaches to solving. As such they work poorly as crowd games even in a simple form, as no guarantee can be made that all players are able to solve a puzzle equally.

3.1.3 Sum up of the State Of The Art review

Through the State Of The Art review several crowd interaction projects were investigated in terms of two things:

- *Technical implementation*: that is, HOW the system was made, which technology was used to capture the intentions and interactions of the crowd and how.

- *Venue and use*: that is, WHERE and for WHAT was the specific system used, projects included both crowd games and other activities such as interactions with music performances, voting mechanics and others.

Generally what is most interesting in relation to designing a crowd gaming system, is the technical implementation. However the combination of the two areas gives valuable insight into what is possible within the boundaries of a venue and or event, and its affordances and limitations.

3.1.3.1 Technical implementations

Generally the State Of The Art review, revealed four main categories of technical implementations, with smaller variations within each category. The categories are summed up here and evaluated in relation to crowd gaming as well as a child audience.

Computer vision based systems:
The majority of projects, investigated in the State Of The Art Review were based on some form of Computer Vision technology, these can again be divided into two sub categories:

- *Object segmentation:* As used in the *Cinematrix* system, The *Squidball* system, the *Glimmer* system as well as several of the examples given by *Maynez-Aminzade, Pausch* and *Seitz* in their *Techniques for interactive audience Participation* article.
  Using this technique, certain objects are segmented from a video image, typically using color segmentation, and features of these objects, are used as input for the system. These features can be anything from the number of objects detected to their position, relative position, amount of movement etc.

- *Template matching:* As used in the *leaning* example by *Maynez-Aminzade, Pausch* and *Seitz, in which* several templates were recorded before the actual interaction, and used in a "closest match" scenario. The technique was also in part used in the *BannerBattle* example, where each new video image was compared to the previous, resulting in a measure of "overall movement".

Using computer vision techniques require limited hardware and prerequisites of the interacting audience. The interaction can be partaken in by anyone, regardless of technical skills and knowledge, and in cases where specific objects were handed to the audience for segmentation, these were typically of a very simple and affordable nature (like colored reflective paddles, glow sticks or laser pointers).
Together these advantages makes a computer vision based system ideal for use in this project.

Audio based systems:
Whether coupled with a computer vision based system, like in the case of e.g. *BannerBattle,* or used as a standalone technique as was the case with the *Cheering meter*, this technique is easily the most simple of the represented systems. In terms of hardware, requiring only a few microphones coupled with some audio analyzing software, and from the audience requiring nothing but the ability to make noise. As such, an audio based system is also viable for use in this project, however the simplicity of these systems make them rather limiting in the possibilities for interaction, typically reduced to a single or two numerical measures of e.g. amplitude or volume and/or pitch.

Phone based systems:
Easily the most complex in terms of the amount of control they can offer a user, are the phone based systems. From the *MobiLenin* project that uses regular personal cell phones, and limits the interaction to a simple voting system, *SpaceBugz!* uses a smartphone application to allow users to both steer, aim and shoot with an onscreen avatar. Finally the composition project presented by *Roberts* and *Höllener*, allowed individual audience members to control live sound parameters of a musical composition by following the motions of the conductor on individual sliders in an application on their smartphone.
Through the use of mobile phones, each user can be given unique information (like the winner of a lottery draw in *MobiLenin* or the score, appearance etc. in *SpaceBugz!*). This allows for more complex rules and gameplay, but also shifts focus from collective efforts towards individual performance. Another issue with a phone based technique in relation to this project, is the requirement of all users to have a phone that the system is compatible with. While many children today might have access to a phone, this is not something that can be guaranteed, and for this reason phone based techniques are not viable in relation to this project.

Gadget based systems:
The last category are the gadget based system, here meaning specifically designed hardware components to facilitate crowd interaction. Reviewed projects included both wearables (necklaces, watches etc.), designed to pick up user movement. As well as low-cost give-away wireless sensors.
Similar to the audio based systems, a specifically designed small piece of hardware could be used in this project. However, as with audio based systems, these allow only for very simple actions to be recorded (like sudden jerks of motion), and therefore typically also only provide a single or two measures (like the amount of movement and rhythm). Move over it would require both time and advanced technical knowhow within the field of electronics to produce a batch of units of sufficient quality, for these reasons this technique is discarded.

3.1.4 Sum up of children and play

Many of the theories and definitions investigated in the initial view on children and play (see section 2.2, page 11), seems primarily concerned with WHY children play, rather than HOW. Theories suggest how children are able to handle and express complex emotions, through play behavior, and how play is an important factor in the development and maturing of children. However theories by *Vygotsky* and *Piaget*, mention an important aspect that might prove useful in relation to designing for children s play. That is the capability of children to separate the meaning of an object from the object itself, through this *pretend play* behavior s simple stick can become a sword or a magic wand, it is reasonable to assume that such imagination can be designed for, and that any object that may be required for interaction within a crowd gaming system, may undergo similar transformations.

3.1.4.1 Sum up of stages and levels of play

What is certainly of interest in relation to this project, are the different forms, or levels and stages that play can take. Looking at the 6 stages of play presented by *Mildred Parten,* these relate primarily to the sociological aspects of play, and it is clear that not all stages are relevant in relation to a crowd gaming experience. As *Parten* states, the social interactions of children develop through age, allowing for more and more stages of play.

While the first 5 stages of play does not have children actively engaging in play together as a group, the last stage, *Collaborative play,* can be directly related to the theories of group and crowd psychology reviewed earlier in the pre analysis (see section 2.1, page 9), doing so it becomes clear that several concepts recur within the two.

In collaborative play, the play activity is governed by rules, there is a shared understanding of the playing activity and of being part of it, participating members have individual roles and depend on each other for the play to function, also this sort of playing activity involves large amounts of social interaction. Based on these observations, it seems reasonable to design a Crowd gaming system to accommodate for a Collaborative playing experience as it is defined by *Parten.* At the same time, all of the levels of play defined by *Piaget,* namely *Functional or practice play*, *Symbolic/pretend/make-believe play* and *Play/Games with rules*, suit the crowd gaming scenario, and as such the system may be designed to accommodate for some or all of these at once.

3.2 Final Problem statement

Having established the primary focus areas for designing a crowd gaming framework for children, the pre-analysis has established a ground frame for this project to build upon.
The most desirable technique to be used has been identified as being based on Computer Vision and Collaborative play has been determined to be the most relevant playing activity to fit the framework.
Moreover, for testing a finished framework, different game genres have been evaluated in relation to their relevance for crowd gaming. Together the pre analysis provides enough background knowledge to device the following Final Problem statement:

"How can a Computer Vision based crowd gaming framework, be designed to feature interaction schemes for games, that can support children in collaborative play?"

More specifically, the framework will be built in and using the Unity3D game engine with a Computer Vision Library plugin. Computer Vision techniques, will focus on color segmentation, as this can be implemented with a minimum of custom designed hardware devises.

To answer the Final Problem Statement, more analysis of Computer Vision and color segmentation techniques. Analysis on how to best conduct tests with a child target group will also be included to provide the basis for a later experiment design.
Based on this analysis, and the theory already established through the pre-analysis, a crowd gaming framework will then be designed and a prototype of it developed.
Several small games will also be designed and prototypes developed, and these will finally be tested in scenarios using a group of children, in an attempt to answer the Final Problem Statement.

4 Analysis

4.1 Computer vision techniques

Being the primary technology used to "sense" the participating audience and their interactions, Computer Vision as a field is of great interest in this project. Specifically, techniques relating to color segmentation, noise reduction, and object tracking are investigated here, to form the basis for design and implementation of a later framework prototype. The section opens with quickly covering some basic features of camera equipment in relation to computer vision, as this has importance for choosing the equipment best suited for implementing the system.

4.1.1 Camera equipment

Two of the main interesting features of a camera, or *image acquisition system* are its lens and sensor.

4.1.1.1 The camera optics

The optics of a camera consists basically of a barrier with a small hole in it known as the *aperture*, and a *lens*, which is basically a piece of "curved" glass, which focuses light rays that pass through the aperture onto the sensor of the camera.
A number of concepts are related to the camera optics, and the lens position in relation to the photo sensor. In relation to capturing a crowd for image processing purposes, these must be taken into consideration.

Focal length (f):
The distance from the optical center (O) to the point Focal point (F) where the lens makes incoming light rays intersect is called the focal length (f). As the focal length is increased so is the size of the captured object in the resulting image, this concept is known as optical zoom. (See Figure 12)

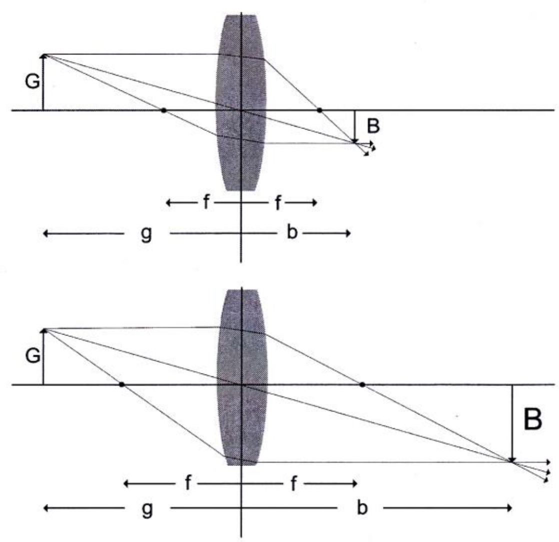


Figure 12: Different Focal lengths f, resulting in different sizes of the captured object B (Moeslund, 2009)

Focus

If the focal length f, remain constant, and instead the distance (g) from the captured object to the lens is changed, the object in the resulting image will also change size.
However, as this distance changes, so does the point where incoming light rays intersect, which in turn means that the distance (b) from the sensor to the lens also has to be changed slightly. The result of not changing b, is an image that appears *out of focus,* in practice this means that instead of only receiving light rays from a single point, each pixel will receive light rays from several different points at once.

Depth Of Field (DOF)

As can be seen in Figure 13, the captured object can be moved a distance g_r closer to, or a distance g_l further away from the camera lens, and still remain in focus. This concept is known as the Depth Of Field of the camera lens and it determines the depth of a motive that can be in focus at any given time. The DOF depends on the focal length f of the lens, as well as the size of the hole through which incoming light has to pass (the aperture).
Depending on what sort of features must be extracted from the motive, and whether these are all generally situated in one plane (like e.g. laser dots or shadows on a large screen) or vary along the depth of the motive as well (like e.g. the reflective paddles held by a cinema crowd), a greater or smaller DOF is required.

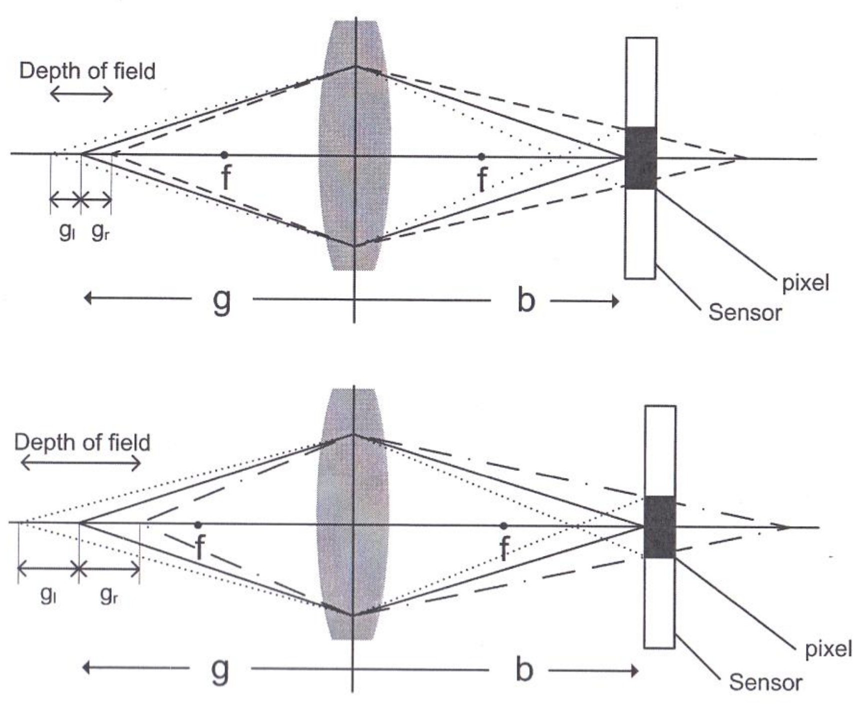


Figure 13: The depth of field of two different focal lengths (Moeslund, 2009)

Field Of View (FOV):
Perhaps the most interesting feature of the lens in relation to capturing a crowd, is the Field Of View. The FOV determines the angle with which the camera "sees". Especially in locations where space can be an issue, and the camera cannot simply be placed further away from the crowd, a camera lens with a large FOV may be desirable, to fit everyone in frame.
The FOV changes as the focal length f is changed, but as illustrated in Figure 14 also depends on the physical size of the cameras photo sensor.

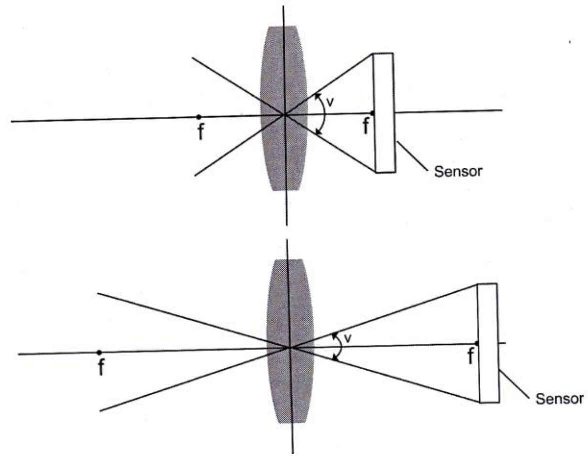


Figure 14: The field of view of two cameras with different focal lengths

4.1.1.2 The camera photo sensor

While the lens of the camera is equivalent to the actual eye or iris of the human visual system, the photo sensor is equivalent to the photoreceptor nerves located in the back of the eye socket. Like the human eye focuses light onto the nerve endings, a camera lens focuses light or energy onto a square area or grid of photo receptor cells, most interesting in this respect is the *size* and *resolution* of the sensor. The size is self-explanatory, and of course is limited by the housing/case in which it has to fit.

Camera resolution:
The resolution of the camera sensor is the number of receptor cells or pixels in the sensor, and determines the number of independent measures of light made by the camera, which will eventually be represented by pixels in each video image. In relation to computer vision, there is a balance between a resolution high enough to capture the desired features of the motive, and a resolution low enough that processing all of the individual pixels each frame won't be to computationally expensive, resulting in too low an end frame rate. Alternatively a high resolution image can always be down sampled, for instance by representing 4 pixels by a single pixel with a values calculated as the means of the 4 pixels it will replace. In principle a low resolution image can also be "up sampled" to a higher resolution, but no information is gained by doing so.

4.1.2 Color segmentation

This subsection will cover the different areas of and techniques within computer vision, relevant in relation to color segmentation.

4.1.2.1 Color spaces

Colors in a camera are based on the same principles as the human eye. While the human eye consists of different nerve-cells that act as photoreceptors, the so called *Rods* and *Cones*, a camera uses one or several sensors to measure the incoming light in different wavelengths, representing colors (Moeslund, 2009). Once captured by the camera, an image can be represented in a number of different ways via so called color spaces, each has its advantages and disadvantages.

RGB color space:
The most common computer representation is the so called *Red, Green, Blue Color space (RGB)*, exactly as its name suggests, the RGB color space represents an image through its Red, Green and Blue components. That means that each pixel has a 3-dimensional vector of values, one for each color (see Figure 15). Typically these values are represented as 8bit integers, meaning that each one has a value representing the intensity of that color, ranging from 0 to 255. Combining all 3 colors using additive color theory means that each pixel can assume 256^3 = 16.777.216 different colors (Moeslund, 2009).

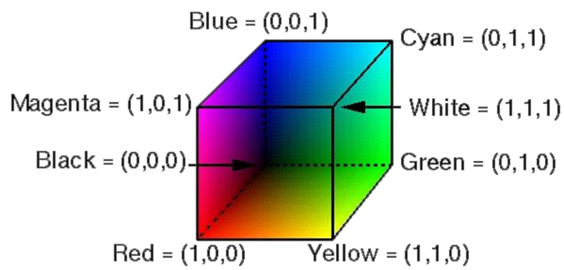


Figure 15: The RGB color cube spanned by the Red, Blue and Green color vectors. Each corner of the cube is a "pure color"; Black, Red, Blue, Green, Magenta, Cyan, Yellow. While the diagonal between the points (0,0,0) and (1,1,1) is the grey vector (Foley, Dam, Feiner, & Hughes, 1990)

HSI color space:
HSI is short for *Hue, Saturation, Intensity,* which is closer to the human way of perceiving colors than the RGB representation.
Hue is the dominant wavelength of light represented in the HSI color space, that is, it is the "pure" color of the perceived light. Saturation is the purity of the hue, how much white light is mixed into the color.
When representing colors through the HSI color space, hue is represented by an angle (0-360 or 0-180 degrees depending on software), while the

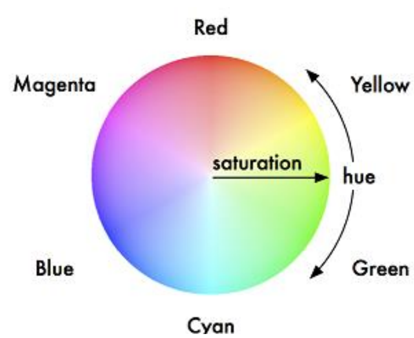


Figure 16: The HSI color circle, with hue represented by the angle around the circle starting and ending with a pure red color. The center of the circle is a shade of grey.

saturation and Intensity typically are represented by integer values ranging 0-255.

HSV color space

The HSV color space is short for *Hue, Saturation*, and *Value*. The HSV color space may be regarded as an approximation of the HIS color space, but much simpler to calculate. While this holds true, it is important to note that HSV is not defined as an approximation of HIS, rather it is defined from an artists point of view. When an artist mixes paint he or she chooses a pure color and lightens it with white or darkens it with black.
The pure color corresponds to Hue, Increasing the brightness of the color by adding white, corresponds to lowering the saturation. While increasing the amount of black corresponds to lowering the intensity of the Red Green and Blue channel. The conversion from RGB to HSV color space is defined in

$$
\begin{aligned}
H &= \begin{cases} \frac{G-B}{V-\min\{R,G,B\}} \cdot 60^\circ, & \text{if V=R and G}\geq\text{B;} \\ (\frac{B-R}{V-\min\{R,G,B\}} + 2) \cdot 60^\circ, & \text{if G=V;} \\ (\frac{R-G}{V-\min\{R,G,B\}} + 4) \cdot 60^\circ, & \text{if B=V;} \\ (\frac{R-B}{V-\min\{R,G,B\}} + 5) \cdot 60^\circ, & \text{if V=R and G} < \text{B.} \end{cases} & H \in [0^\circ, 360^\circ[\\
S &= \frac{V - \min\{R,G,B\}}{V} & S \in [0,1] \\
V &= \max\{R,G,B\} & V \in [0,255]
\end{aligned}
$$

Equation 1

$$
\begin{aligned}
H &= \begin{cases} \frac{G-B}{V-\min\{R,G,B\}} \cdot 60^\circ, & \text{if V=R and G}\geq\text{B;} \\ (\frac{B-R}{V-\min\{R,G,B\}} + 2) \cdot 60^\circ, & \text{if G=V;} \\ (\frac{R-G}{V-\min\{R,G,B\}} + 4) \cdot 60^\circ, & \text{if B=V;} \\ (\frac{R-B}{V-\min\{R,G,B\}} + 5) \cdot 60^\circ, & \text{if V=R and G} < \text{B.} \end{cases} & H \in [0^\circ, 360^\circ[\\
S &= \frac{V - \min\{R,G,B\}}{V} & S \in [0,1] \\
V &= \max\{R,G,B\} & V \in [0,255]
\end{aligned}
$$

Equation 1: The conversion from RGB to HSV, where min{R,G,B} and max{R,G,B} are the smallest and biggest of the R, G, and B values respectively (Moeslund, 2009)

4.1.3 Color thresholding

Color thresholding is a powerful approach to segmenting out objects in a scene.
Typically, color thresholding function like the band pass filters known from other areas of signal processing. Regardless of the color space (RGB, HSV, HSI etc.) each pixel in the image has each of its 3 components compared to a minimum and maximum threshold value. If all 3 values lie within the min. and max. range the pixel color is set to while (foreground pixel) otherwise it is set to black (background pixel).
To reduce computation costs in further processing of the image, these pixels can be set in a

separate binary output image, consisting solely of zeroes and ones.
Effectively, this sort of color thresholding algorithm corresponds to defining a small box within the RGB color cube, and determining if the current pixel lies somewhere within this box or outside it.
One problem with color thresholding I that it is very sensitive to changes in illumination in the captured scene. As the illumination changes, so does the color values captured by the camera. To account for this, the range between the minimum and maximum threshold values have to be extended (the box made larger) to account for all possible color values of the desired object. This of course means that the algorithm becomes more prone to errors, and the risk of including objects that are not of interest becomes greater. In the very worst case, the cube will be as large as the entire RGB cube, and all objects in the scene will be included in the segmentation.
The solution to this problem is to first convert the image into a color space where color and intensity are separated. And then do thresholding only on the colors (e.g. HS), disregarding their intensity. The thresholds can in this case be kept more "tight", reducing the risk of false classification. (See Figure 17)

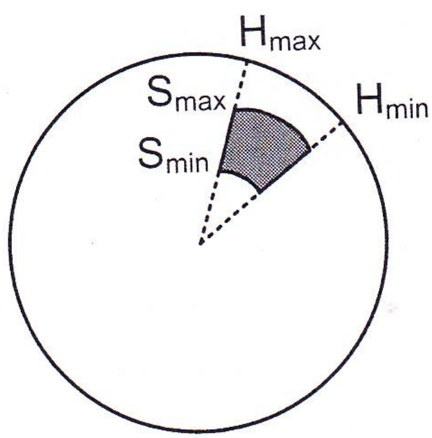


Figure 17: Example of threshold values in a HS color space,
the grey area within the color space is where segmented objects lie (Moeslund, 2009).

Some thresholding on the intensity is however often a good idea, as it is generally very hard to distinguish colors with very low or very high intensities, as these will be interpreted as very close to black or white respectively, containing very little color (Moeslund, 2009).

4.1.4 Morphology

Once an image has gone through a color threshold, you are left with what can essentially be regarded as a binary image of "ones" and "zeroes". However, as mentioned above, color thresholding is prone to errors due to the need to keep the min. and max. range large to accommodate for different illumination. The result is over segmentation in some areas, while typically even with a very "generous" minimum threshold value, some areas are still under segmented. This means that while the color thresholded image has the objects of interest

segmented as white pixels, it will typically also have a number of white pixels which do not represent objects of interest, these pixels are regarded as noise.
Furthermore the objects of interest may have holes in them as a result of under segmentation.
All of these problems can be dealt with using different algorithms of what is known as *mathematical morphology.*

4.1.4.1 Kernels or structuring elements

As with other neighborhood processing algorithms within the field of image processing, morphology algorithms operate by applying what is known as a kernel, to each pixel of the image. In morphology these kernels are known as *structuring elements*.
Kernels are essentially a certain sized grid of pixels with a value of either 1 or 0.
kernels can be designed as one pleases, but typically the pattern of ones form a box or a disc.
Figure 18 shows some examples of different kernels.

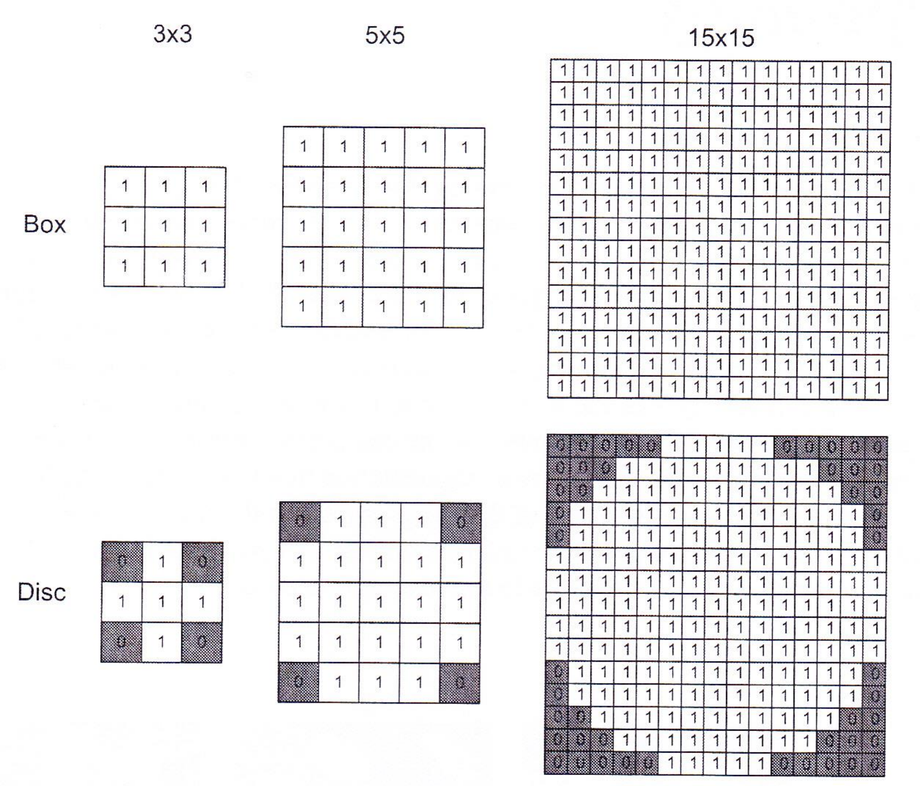


Figure 18: illustration of 2 different structuring elements or kernels (a square and a disc) at three different sizes (3x3, 5x5 and 15x15 pixels)

4.1.4.2 Hit and Fit

While other image processing algorithms that use kernels typically operates using addition and multiplication methods, morphology is based on a concept known as *Hit and Fit*.
The structuring element is centered on top of the pixel in focus (a white pixel) and it is this pixels

resulting value that will be determined based on the structuring element and the morphology operation.

Hit
For each of the 1's in the structuring element the corresponding (underlying) pixel in the image is investigated. If just one of these also has a value of 1, the structuring element is said to have *hit* the image at this particular position, and the pixel in focus (the pixel the structuring element is centered on) is set to 1.

Fit
For each of the 1's in the structuring element the corresponding (underlying) pixel in the image is investigated. If all of these also have a value of 1, the structuring element is said to *fit* the image at this particular position, and the pixel in focus (the pixel the structuring element is centered on) is set to 1.

4.1.4.3 Dilation

Applying a Hit operation to an entire image, is known as the morphology operation called *dilation.* It is called so because it in essence dilates the white pixels of an image, increasing the size of white areas.
Effectively dilation increases the size of segmented object, small gaps in the objects are closed and segmented objects that are close to each other are merged together.
Dilation will however not only increase the size of objects of interest but also the size of noise objects. Fortunately there are morphology operations to deal with this.

4.1.4.4 Erosion

Applying a Fit operation is known as the morphology operation called erode. While dilation enlarged white areas of the image, erosion "erodes" them, tearing away at all edges of white areas. The result is that segmented objects become smaller and more round (as small branches, spikes etc. from the object are eroded), small noisy object may be removed entirely by an erosion. In Figure 20 and Figure 19 a dilation and an erosion operation are compared.

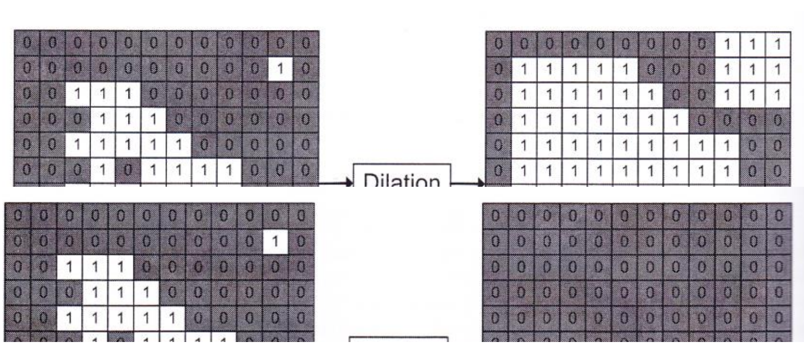


Figure 19: Example of a dilation operation on a binary image - holes are closed and objects grow and melt together (Moeslund, 2009)

Figure 20: Example of an erosion operation on a binary image - object schrink in size or are removed intirely, also any small details along edges disappear (Moeslund, 2009)

4.1.4.5 Compound operations

Combining erosion and dilation in different ways results in strong image processing algorithms known as *compound operations*, the two that are of interest for this project are described here

Closing:
Closing deals with the problems associated with dilation, e.g. the increased size of objects and noise, and the melting together of objects.
Applying an erode operation after a dilation results in a closing operation. Objects that have been melted together are separated again, and objects are returned to roughly their original size, however holes in the objects remain patched.

Opening:
Opening deals with the problems of objects becoming smaller and fractured into smaller individual parts when erosion is used to remove noisy pixels in the image and to "round off" objects. Applying a dilation operation to an image after an erosion results in what is known as an opening. The object is returned to roughly its original size, and fractured objects are melted back together, while they remain smooth. Noise also does not reappear, but will remain removed from the image.

Combining compound operations:
In some cases it is necessary to combine opening and closing operations, e.g. in cases where there are both holes in the objects of interest, as well as noisy pixels around them.
One thing to note is that the structuring element of the closing and opening operations need not be the same, and the best results are often reached by having e.g. one of them be bigger or in another shape that the other.

4.2 Designing crowd activities

Based on the tests of the games described by *Maynes-Aminzade, Pausch,* and *Seitz* (see section 2.5.3, page 26) a number of guidelines for developing similar systems is presented in their paper (Maynes-Aminzade, Pausch, & Seitz, 2002).
These guidelines are included here as they serve as valuable information, in relation to selecting activities to work as input methods for this particular project, as well as design fitting content to match both the requirements of the child audience as well as the affordances and limitations of the input. The guidelines are divided into those relating to System design, those relating to Game Design, and those relating to Social Factors and as such will be presented here likewise.

4.2.1 Guidelines relating to System Design

- *Focus on the activity not the technology***;** the paper argues that while an audience is initially amazed with the technology behind the system, this amazement quickly subsides if the activity is not entertaining.

- *You do not need to sense every audience member*; this statement is based primarily on their template machining motion detection approach. The paper argues that what matters is what the audience thinks is going on with the system, not what is actually going on. They found that even though an audience member is outside of the cameras field of view, he or she will participate fully in the experience, and enjoy it just as much as those who are inside, simply because they do not know that their actions are not actually recorded by the system.

- *Make the control mechanism;* the paper here argues that even though the underlying technology need not be exposed, it is important to make the control mechanisms obvious and make the audience members understand how their actions affect the game activity. Audience members will not continue to participate in a game activity if there is not an immediate indication that the overall collective behavior of the crowd are affecting the game-play.

4.2.2 Guidelines relating to Game Design

- *Vary the pacing of the activity*; the paper concludes that the leaning and beach ball batting games worked best when they included a pacing that alternates between moments of intense activity and moments of relaxation. For instance Pong works better than the racing game, as it gives the participating audience some "deadlines" that lets them succeed (making the ball hit the paddle), and rest periods to cheer, applaud themselves, and prepare for the next moment of tension

- *Ramp up the difficulty*; this is a classic game design principle, and the paper just underlines how this also works for crowd games. They found that their games did not require any explicit tutorial if they were presented properly and started at a low enough difficulty level. By starting slow and gradually ramping up the complexity of the game, they avoid tedious training phases that work even worse in a crowd than with single players (where the player can skip a tutorial once he or she understands the rules being explained) as parts of the group will understand the concepts explained faster than others.

4.2.3 Guidelines relating to Social Factors

- *Play to the emotional sensibilities of the crowd*; the paper here goes to state that social involvement with the system and its activities are more important than technological involvement. As an example they mention how their laser pointer games were often played with audiences where not nearly everyone got a laser pointer, and how these were the best shows, since they promoted and encouraged the rest of the crowd to shout, give advice and cheer for the players holding the lasers. Similarly, the beach ball game had no trouble engaging the whole crowd via what they call "The lottery effect". Simply meaning that even though only a few members of the audience were in contact with the ball at any given time, the feeling of “I might be next” and the cheering and booing of other players performance, added greatly to the experience.

- *Facilitate corporation between audience members***;** the games described in the paper were generally more engaging when they fostered a sense of camaraderie amongst players. While Whack-a-mole has each player playing more or less for himself, a game of connect the dots works much better from a social perspective, as it requires each player to focus their laser pointer on a different dot, which again requires communication and corporation between players.

4.3 Testing with a child target group

In the field of usability testing with children three common target age ranges exist: *preschool-aged children* (2 to 5 years), *elementary-school–aged children* (6 to 10 years), and *middle-school-aged children* (11 to 14 years). These age divisions are however arbitrary, and many behaviors within them overlap. A general consensus is also that when it comes to the usability of computer products, most children younger than 2 ½ years of age are not proficient enough with standard input devices (e.g., mouse or keyboard) to interact with the technology in a meaningful way, while children older than 14 years of age will likely behave as adults in a testing situation and should be treated accordingly (Hanna, Risden, & Alexander, 1997).

Hanna, Risden, and Alexander states that children in the middle-school age are relatively easy to include in software usability testing. Their experience in school makes them ready to follow directions from an adult, and they are generally not self-conscious about being observed as participate in a test. They are also willing to answer questions and try new things, and will develop more sophisticated ways of explaining wht they see and do.

4.3.1 Guidelines for testing with children

Hanna, Risden, and Alexander also present a number of guidelines for test design in relation to children. Although the guidelines are related to traditional software usability testing (like choosing preferred designs or completing a series of tasks), a number of them are useful in relation to other game and system design evaluations as well. The slightly modified ones included here have been deemed to extend beyond traditional usability testing and provide valuable methods for a later experiment design involving a Crowd Gaming system.

- *Make the test environment more child friendly:* but don’t go too far. Strike a balance between an adult-oriented test environment and an inviting play space that may distract children from the task.

- *Use laboratory equipment effectively but unobtrusively:* make sure to place microphones close to children to pick up their soft voices. For interviews, avoid furniture arrangements that face children directly toward a video camera.

- *Establish a relationship:* by engaging in some small talk to find out more about one another. Children are often happy to talk about their birthdays, their favorite computer games, or their favorite subjects or sports at school.
- *Have a script:* for introducing children to the testing situation.
- *Motivate children:* by emphasizing the importance of their role. For example, tell them that you have forgotten what it is like to be a child, and that you need their help to make a good product for children all around the world.
- *Set expectations appropriately*: for what the children will be doing. Many children will expect to see a finished product and may be disappointed when they are presented with an unfinished prototype. Explain to them why it is important to get their feedback at the current state.
- *Observe:* how much children like a product by observing signs of engagement such as smiles and laughs or signs of disengagement such as frowns, sighs, yawns, or turning away from the task. These behavioral signs are much more reliable than children's responses to questions about whether or not they like something. Children are eager to please adults, and may tell you they like your product just to make you happy (Hanna, Risden, & Alexander, 1997)

5 Design

This chapter will first present requirements for the design of the different aspects of the crowd gaming framework, based on the problem statement and topics covered in the analysis and pre-analysis chapters. The chapter will at the same time present the chosen design solution to each of these aspects and their requirements. These include both the choice, and combination of hardware and software, as well as the choice of which parameters of the participating crowd, should be measurable by the system.
Finally the chapter moves on to cover the game design choices behind 3 small games used for testing the framework, and how these are each designed to utilize and evaluate the capabilities of the framework in different ways.

5.1 Framework design

The design of the crowd gaming framework is centered around 3 main aspects; The choice and setup of camera hardware, the selected combination of image processing algorithms and objects to be captured, and finally what sort of interactions this combination make available for game input in Unity3d. The general setup of the framework is illustrated in **Error! Reference**

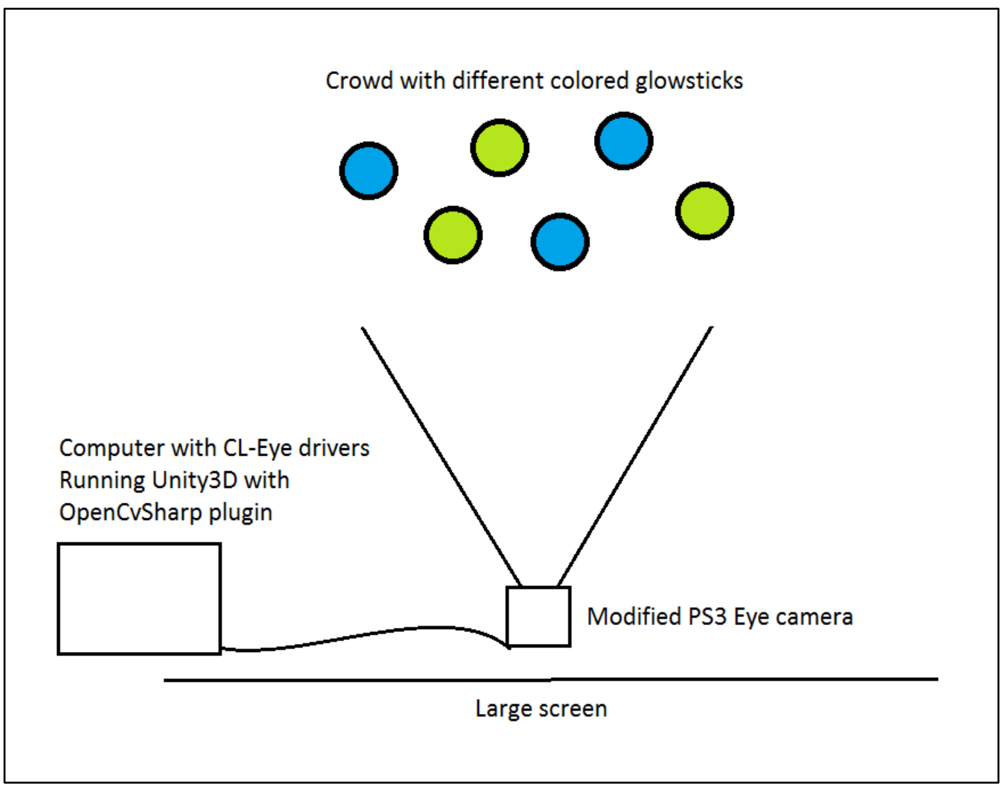


Figure 21: Illustration of the overall framework setup.

source not found..

In the following sub sections each aspect is explained in terms of established requirements and chosen solution.

5.1.1 Camera Hardware

The beginning of any computer-vision system, starts with the choice of camera hardware. It is the camera that will deliver images to the system, and hence it forms the basis of everything that will since happen to these images in the system.
As investigated in section 4.1.1 "*Camera equipment*", quite a few different characteristics are related to cameras and for a crowd gaming framework, some requirements of the camera can be listed:

- The camera needs to be of a wide enough FOV to capture the entire crowd at an acceptable distance

- The camera needs to be of a high enough resolution be able to provide enough detail to capture and segment the desired object(s)

- The camera needs to be of a high enough frame rate to make interactions and response from the system smooth

- The camera needs to be able to provide this frame rate, even in low or dimmed light environments

- The camera needs to provide full control over all image sensor and post processing settings, including white balance, gain, exposure etc.

5.1.1.1 The Playstation 3 Eye camera

In the following sub section, the Playstation 3 Eye camera, that was chosen for this project is presented in terms of its features, specifications and the modifications made to suit the project.

Features:
The Playstation 3 Eye camera (PS3 Eye) is a digital camera device similar to a webcam originally intended for use with computer vision based gaming applications for the Playstation 3 console. The camera is able to capture both video of a resolution of 320X240 pixels, at up to 120 Hz (frames pr. Second), as well as 640x480 resolution at 60 Hz. As a standard the camera features a 56/75 degree FOV turn lens. The image sensor of the camera is designed to have larger pixels than the standard webcam, to make the camera able to *"produce reasonable video quality under the illumination conditions provided by a TV set"* (ThreeSpeech, 2007)
These features makes it immediately ideal for computer vision based applications, compared to standard web cameras within its price range.

The camera also features a build-in-four-capsule microphone array which may also be utilized in connection with a crowd gaming application.

Drivers and settings:
Windows and Mac drivers for the PS3 Eye camera are available through codelaboratories.com, making it possible to use the camera with video chat applications like Skype and MSN Messenger, while at the same time making it possible to access the camera from image processing libraries like OpenCV.
Additionally, codelaboratories provides a Test application, with interface to control camera parameters like Gain, Exposure and white balance (see Figure 22).
These settings are saved directly into the camera driver and applies to other applications as well. This means that settings can be custom adjusted in the CL-Eye Test application and then used in the Crowd gaming framework in Unity3d.

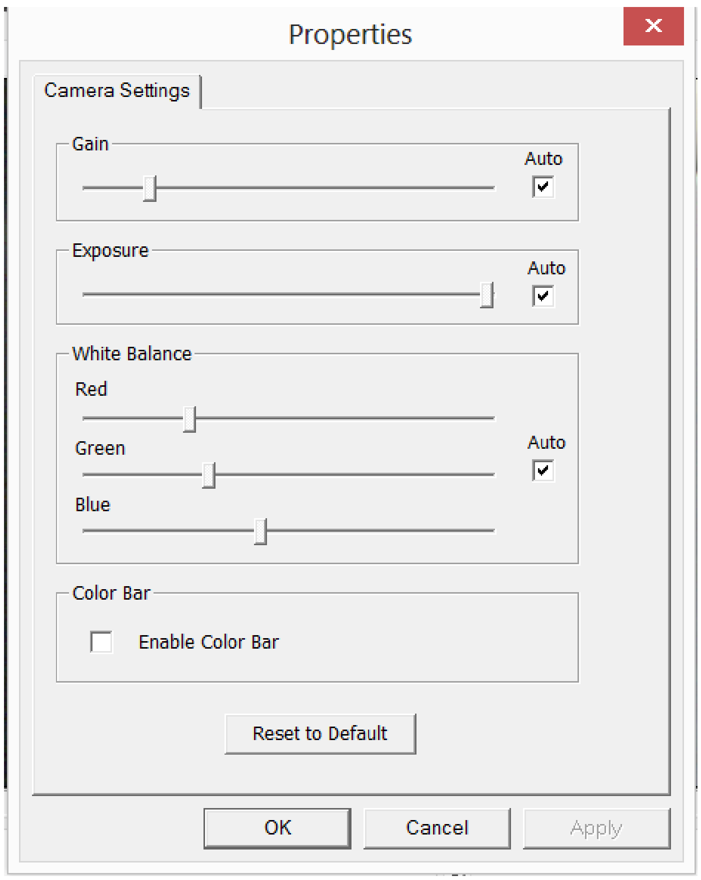


Figure 22: The properties panel of the CL-Eye Test application, in which Gain, Exposure and White balance of the camera can be controlled (codelaboratories.com, 2013)

Modifications to the camera lens:
The PS3 Eye camera uses a standard M12x0.5 S-mount (see for fitting its standard lens. This is the same type of mount which is also found in many types of CCD chipset CCTV surveillance cameras. This means that the lens can fairly easily be replaced by a wide variety of lenses available online. (See Figure 24 and Figure 25)

Figure 24: The M0.5 mm. theaded S-mount that fit inside the PS3 Eye camera

Figure 23: A number of different lenses with different focal lengths and FOV for CCTV surveillance cameras

To get an even wider FOV, making it possible to capture a wider audience without having to increase the distance between the camera and crowd, a PS3 Eye camera was refitted with a 160 degree FOV lens with a focal length of 2.1 mm. (see Figure 26 and Figure 27).
More than doubling the FOV of the camera means a much larger (and especially wider) audience can be captured at relatively short distances.

Figure 25: A traditional PS3 Eye camera with its standard turn lens

Figure 26: Customized PS3 Eye camera, refitted with a 160 degree FOV lens.
Focus is adjustable by turning the lens, thereby adjusting how deep into its mound it is fastened.

5.1.2 Unity3d

Since its initial release in 2005, Unity3d has been an ever growing game development engine and has gone from initially being only an OS X game development tool, to now supporting many different platforms, including Windows, Linux, iOS, Android, Windows phone, Blackberry, Flash, Web browsers, PlayStation 3, Xbox 360 and the WiiU (Unity3d.com, 2013).
Unity is in its basic version free, and its workflow makes rapid prototyping of games and other interactive content fairly easy, this combined with the possibility of including so called Plugin programming libraries for e.g. image processing, makes Unity an obvious implementation choice.

5.1.2.1 OpenCVSharp

Originally launched in 1999 as an Intel research initiative, OpenCV was, and continues to be a strong programming library mainly aimed at real time computer vision applications.
OpenCV is written in C++ and its primary interface is also C++, but interfaces in Java, and MatLab also exists.
Moreover wrappers in other languages such as C# and Ruby have been developed to encourage adaption to wider audiences and uses.
OpenCVSharp is such a wrapper, it wraps the main Libraries of OpenCV into a C# translation ((OpenCvSharp, 2013)
Building the OpenCVSharp Libraries as .dll packages allows them to be imported into Unity3D as a Plugin. This makes common OpenCVSharp functionality available in Unity projects, and as such allows e.g. prototyping of (crowd) games that make use of image processing for interaction purposes.

5.1.3 Glowsticks as captured objects

In line with other successful projects like the Cinematrix, Squidball and Audience/performer interaction project, the crowd gaming framework designed in this project will focus on capturing physical objects or “tokens” held by the children interacting with the framework.
More specifically focus will be put on capturing two different colors of Glowsticks or Chemical lights as they are also know.
Glowsticks consists of a flexible plastic tube with a chemical compound inside, in the center of the tube is another small cylinder made of glass, with another compound inside.
As the glowstick is bend, the inner glass cylinder snaps and shaking the glowstick will now mix the two chemical compounds resulting in a chemical reaction that emits a strong light.

(See
Figure 27)

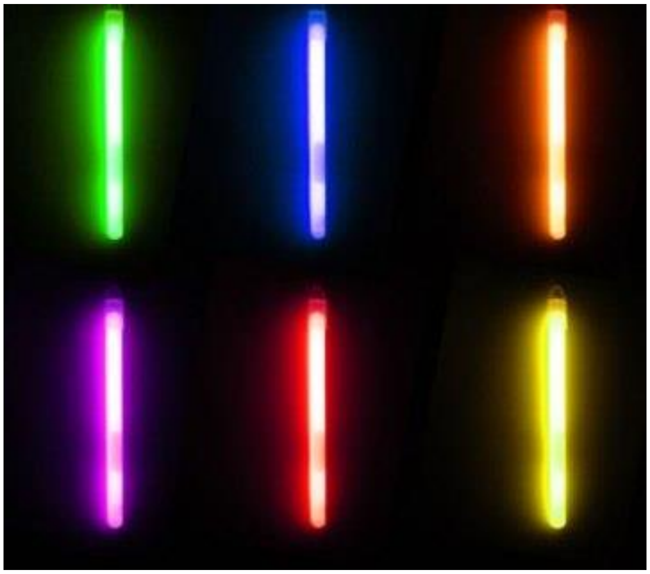


Figure 27: Different colors of activated glow sticks emitting light in the dark

Glowsticks are ideal as tokens to be segmented by color segmentation due to their highly saturated colors and ability to emit light by themselves.
For a crowd gaming framework for children, they are optimal, since they are inexpensive and disposable. They require no additional light sources to be shone upon them like e.g. the reflective paddles used by the cinematrix system, will work very well in dim lighting conditions and their multitude of color options provides for interesting and natural division of the engaged players into separate collaborating or competing groups.
As a starting point, the framework is designed and implemented to be able to segment two different colors of glow sticks, green and blue.

5.1.4 Measureable parameters

Once segmented it must be decided what to actually use as input parameters for a game that uses the framework.
A number of different options are available from individual positions, overall (or mean) positions, positions relative to the last captured frame (optical flow), gesture recognition etc.
While a complete framework should ideally support many or all of these options, the design choice was made to initially include only the most simple of these, namely the overall mean position of green and blue glow sticks respectively.
While a very simple measure reduced to a mere two (x, y) coordinate sets (one for each color) for each frame, a number of interaction possibilities are possible with just this data, as will become apparent in the Game design chapter.

5.1.4.1 Measurable parameters and collaborative play

Using a mean position of all glow sticks of a color inherently forces players to work together in unison to effectively interact with the framework, and as such offers possibilities for collaborative play. At the same time this type of interaction offers possibilities for different kinds of groups dynamics, self-organization into decision maker/leader, follower, onlooker roles etc. Furthermore, once a mean position of the glow sticks is recorded every frame it is a mere matter of subtracting the position from the previous frame from that of the current to detect simple collaborative gestures like waving or swaying from side to side or up and down.

5.1.5 Three interaction paradigms

Based on the parameters measured by the framework (mean x and y positions of green and blue glow sticks), 3 interaction paradigms were designed, each one slightly more complicated than the previous one.

5.1.5.1 Single axis single color interaction

The first and simples interaction is based on just one color of players working together to move glow sticks on just a single axis (in the implemented case the x axis), corresponding to moving horizontally - left/right.

Both colors can of course do this simultaneously, controlling each their own game avatar in some form of competitive or collaborative gameplay scenario, or the system may disregard color differences entirely, thus making everyone control the same avatar simply by collectively moving left or right to adjust the mean position of all glow sticks.

This position is then mapped directly to the game avatar(s), e.g. in a game of *pong* or similar where the mean glowstick position would correspond to the position of a paddle.

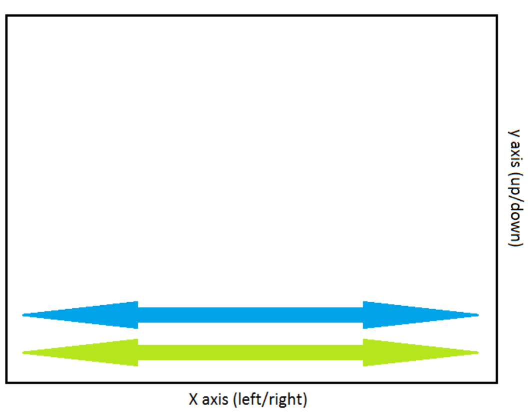


Figure 28: ilustration of the single axis single color interaction paradigm, with the mean position of each color of glow sticks being evaluated only on the x – axis

Single axis dual color interaction A slightly more advanced interaction paradigm has two colors working together to control the same game avatar. As with the first interaction paradigm each color moves only on a single axis. But while the first interaction paradigm has all glow sticks

moving on the same axis, this paradigm has each color moving on its own separate axis (see Figure 30).
Again, mean glowstick positions may be directly mapped to avatar positions (but no on two axis) or they may be evaluated as being to the left/right, or above/below the center of their respective axis, the important note is that the two colors must work together collaborately to control an avatar properly on both axis)

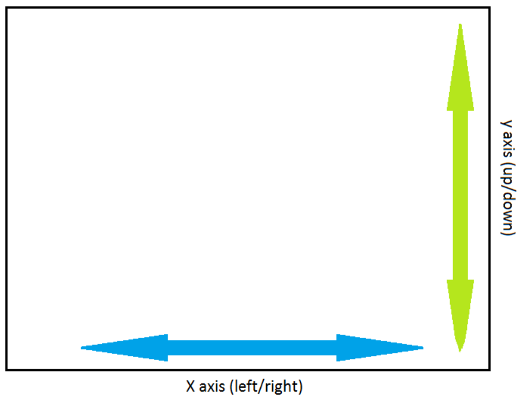


Figure 29: Ilustration of the single axis dual color interaction paradigm, with the mean position of each color of glow sticks being evaluated on each its own axis

5.1.5.2 Dual axis dual color interaction

The third and final interaction paradigm, is also the most complex one.
While the two previous paradigms had each color of glow sticks being evaluated on a single axis, this paradigm has both color of glow sticks being evaluated on both axis simultaneously. This means that players have to move their glow sticks on both the x axis (horizontally) and the y axis (vertically) simultaneously (see Figure 30).
This interaction paradigm allows for competitive gameplay, the colors inbetween, in a number of gameplay scenarios, as each color group can be provided with more complex (two dimensional) interaction possibilities with each their own game avatar.

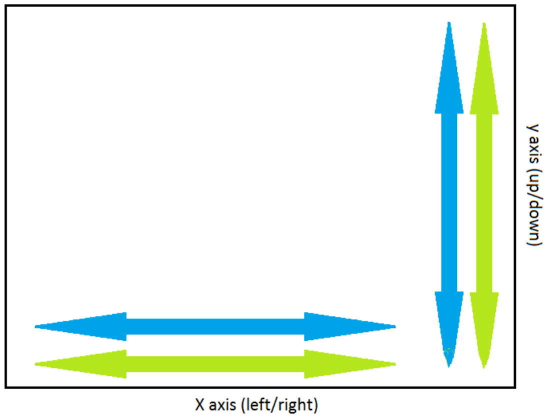


Figure 30: Ilustration of the dual axis dual color interaction paradigm, with the mean position of each color of glow sticks being evaluated on both the x and the y axis simultaneously

5.2 Game design

The following section will contain descriptions of and reasoning behind some of the design choices made in the development of 3 small games used to test and evaluate the crowd gaming framework with the measurable parameters and interaction paradigms that were designed and implemented for the project.
The Game design section contains first a short description on some general design choices that are valid for all of the games, and how these are justified.
This is followed by a presentation of the design of 3 games that each is based on one of the 3 previously presented interaction paradigms.

5.2.1 General design choices

First and foremost it is important to note that none of the 3 games developed are grand game masterpieces, rich in 3d assets, high quality textures, custom sounds and soundtracks.
Rather, they are short games with very basic gameplay and only the most essential mechanics and assets implemented to be able to test the framework. The games are all in the arcade or vehicle simulation genre, as these were analyzed as being best suited for crowd gaming purposes. During the design of these games, focus has been on fulfilling the basic qualities and features that constitute a game with non-trivial gameplay, as investigated in the analysis chapter (see section 2.4.1 - *Game definitions.*)

5.2.2 Designing a Breakout game

The first game is designed to test the framework using the first interaction paradigm ("Single axis single color interaction") *The Breakout game* is a rendition of the classic arcade game *Breakout*, introduced in 1976 by *Atari*.
Breakout was originally directly influenced by the 1972 game *Pong*, and similarly to pong, had the player(s) steer a paddle sideways on the screen, trying to hit a small ball as it bounces

around on the screen edges. What makes Breakout different from pong, is a wall of blocks situated in the top of the screen that the player can “break” by hitting the individual bricks one or several times with the ball, points are gained for breaking blocks.
The bottom of the screen (below the paddle) serves as an end condition for the game, the player will lose by failing to hit the ball and letting it pass below the paddle. (See Figure 31)

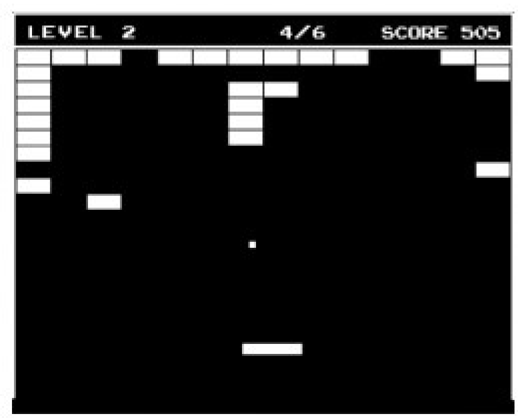


Figure 31: The original 1976 game: Break Out by Atari ((Atari, 1976)

Through the years a number of versions with various gameplay alterations have been published by a number of developers, with the core gameplay however, remaining the same – hitting a ball with a paddle.

5.2.2.1 Alterations to the original design

As with the original design, the Breakout version designed and implemented for this project is also based on the simple mechanics of hitting a ball by moving a paddle left or right on the screen, apart from the paddle being moved collectively by a crowd of players (as described in the previous section on interaction paradigms) this version has a few alterations to the gameplay as well (see Figure 32 on the following page).

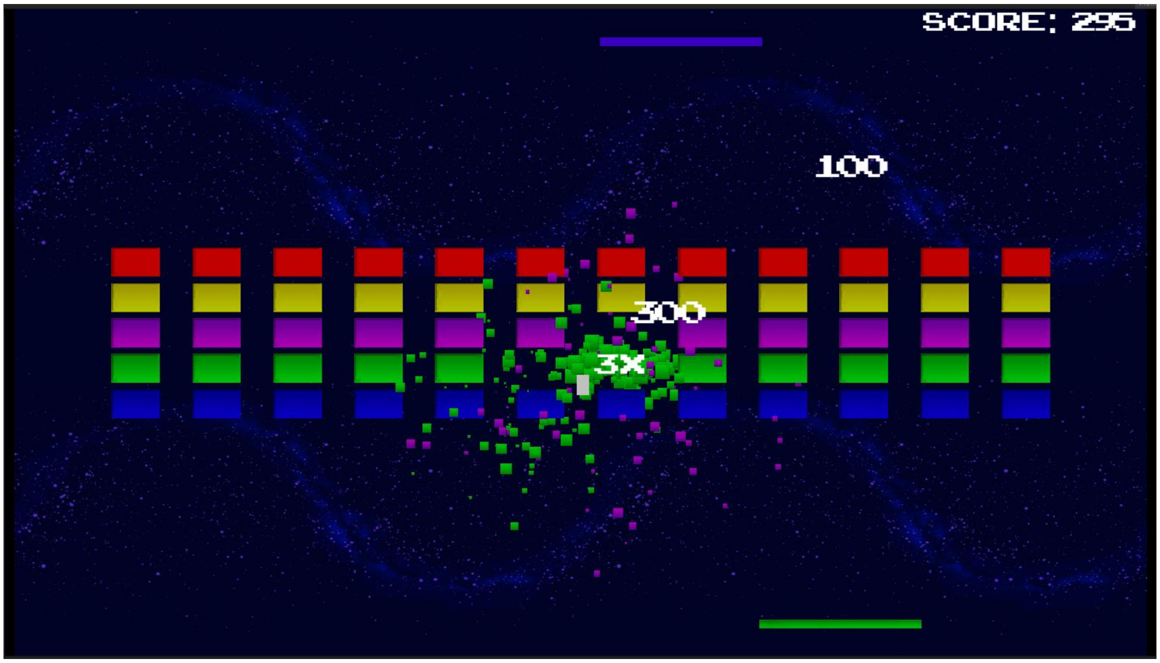


Figure 32: Screenshot of the Breakout game designed for this project, the two paddles of different color are seen in the top and bottom of the screen. The ball is the small grey square in the middle of the screen, in the process of breaking blocks in a cascade of colored pixels

1. Rather than having just one paddle in the bottom of the screen, this version has two paddles, one for each color of glow sticks. Paddles are located in the bottom and top of the screen and individually colored in the same color as the glow sticks that control them.

2. The end condition for this version of the game is located two places (one for each group/color of players), below the bottom paddle and above the top paddle.

3. The wall of blocks are situated in the middle of the screen, with even distance to the two paddles

4. The blocks are colored in a number of colors and various score bonuses or “combos” are designed for hitting many blocks in a row, consecutively hitting blocks of the same color and hitting just one of each color of blocks in a single series.

5. Visual feedback is provided by letting the blocks “explode” in a cascade of pixels, while points and bonus triggers pop up as large written texts before flying up to be added to the score counter in the top right corner.

6. Auditory feedback is provided via different sounds that are regulated in pitch as combos rise.

As a crowd game, this version of Breakout becomes a collective as well as a collaborative effort. While each group has full control over their own respective paddles, it becomes a collaborative effort between the two groups of players (glowstick colors), to break as many blocks as possible. The points gained by each group of players is not recorded separately but are tallied together in the same total score, and each group of players are dependent on the other group to return the ball to their side of the screen.

5.2.3 Designing a Flying game

The second game is designed to test the framework using the second interaction paradigm ("Single axis dual color"). *The Flying game* is in the vehicle simulation genre and is based on a very limited number of game assets and mechanics.
The flying game has players steering an airplane through giant hoops in a cloud filled the sky, while trying to avoid dark clouds. (See Figure 33)

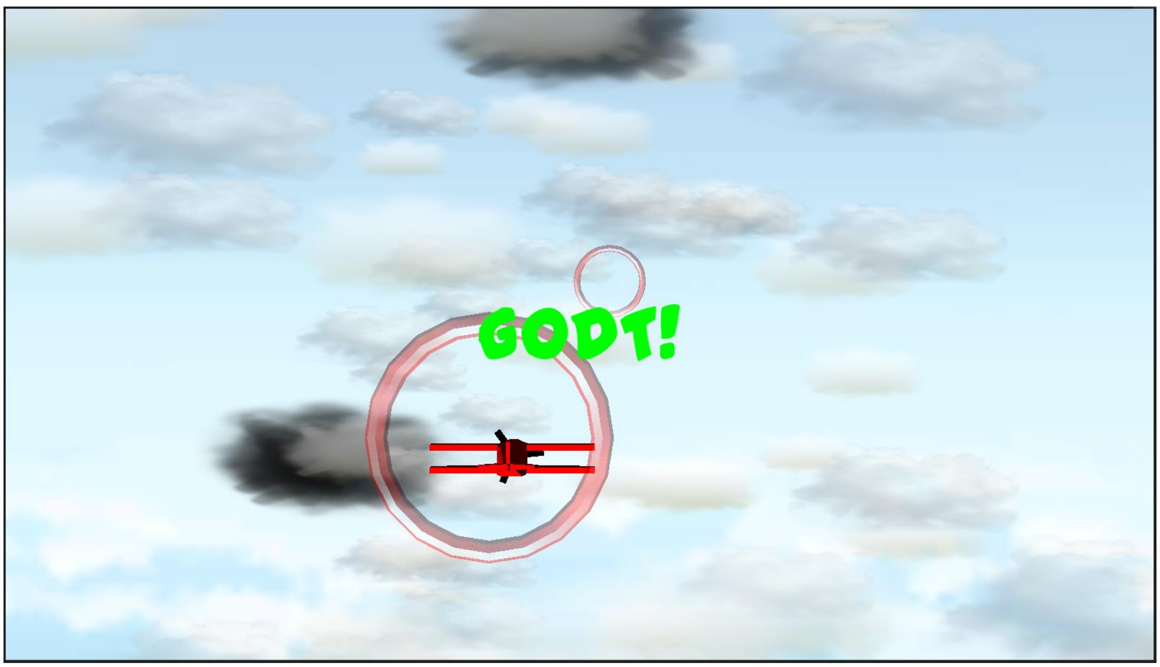


Figure 33: Screenshot of the Flying game designed for this project. As the red airplane passes through hoops in the sky, visual feedback in the form of large green texts appear on the screen along with an auditory "pling" sound.

As with the breakout game described above, as a crowd game the flying game becomes a collective as well as a collaborative effort.
Different from the breakout game where each group of players control their own avatar (paddle), the flying game has the two groups of players collaboratively steering the same plane. One group collectively steers the plane up and down by raising or lowering their glow sticks, while the other group collectively steers the plane left and right by moving from side to side with their glow sticks in front of them. Collaboratively they gain full control of the plane.

5.2.4 Designing a Racing game

The third and final game is designed to test the framework using the third interaction paradigm ("dual axis dual color"). *The Racing game* is also in the vehicle simulation genre and similarly to the Flying game based on a very limited number of game assets and mechanics. The Racing game has players steering small go-karts on a track. The track has both left and right turns, and its sides are blocked by walls, prohibiting the go-karts to drive off it (see Figure 34).

Figure 34: The track for the Racing game as seen from an elevated perspective view in the Unity3d editor

As is apparent from the screenshot in Figure 35 on the following page, the racing game differs from the two other games in a number of ways:

1. The game is played in a competitive split screen mode

2. Each group of players (color of glow sticks) has full control over their own go-kart avatar, and races the other group competitively around the track.

3. As dictated by the "two axis dual color" interaction paradigm, each group of players must move their glow sticks on both the x and the y axis to fully control their avatar.

4. Instead of being directly mapped onto onscreen positions for the avatar, the mean glowstick positions are for this game interpreted as being either in a "dead zone" in the middle of, or to the right/left or top/bottom of the x and y axis respectively

5. The left/right x-axis interpretations, determine if the go-kart goes straight ahead or turns, while the top/bottom y-axis interpretations determine if the car accelerates, brakes or reverses

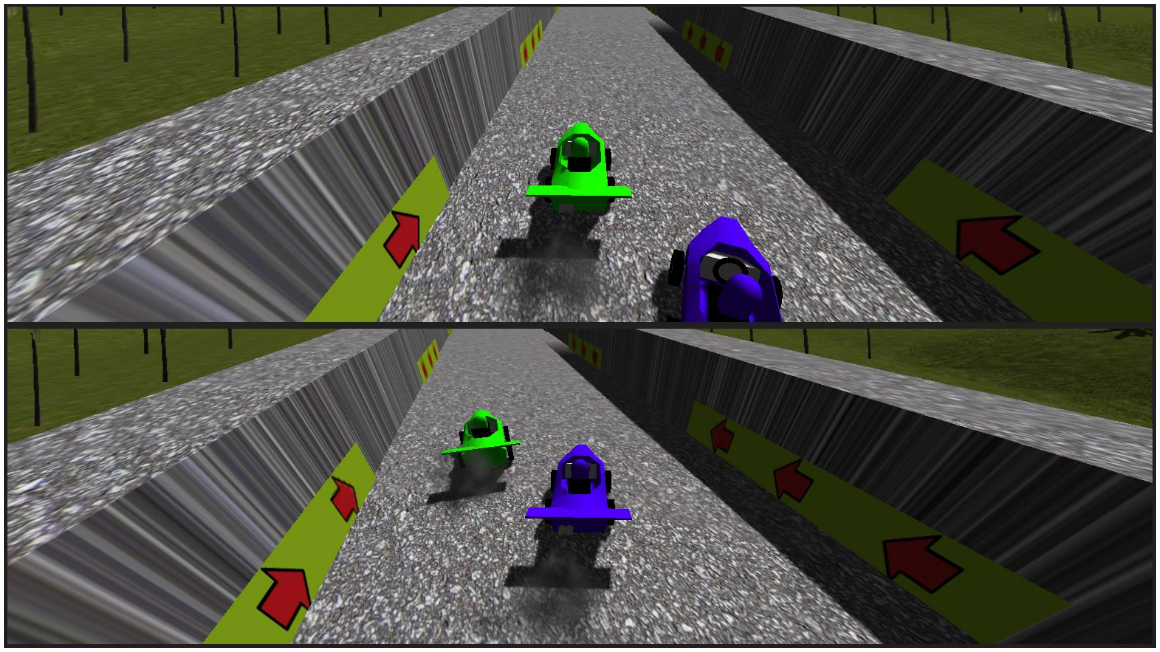


Figure 35: Screenshot from the Racing game, the game is played in split screen with a group of players controlling either the upper green car or the lower blue car around the track.

Different from the other two games, the Racing game also does not keep scores, and does not provide visual feedback in the form of text. Instead it is always apparent who is in the lead, due to the split screen design, and turning the go-kart along with the track and avoiding the walls, provides visual feedback in itself, on how well each group of players are doing

6 Implementation

The following chapter will go into some implementation specific details in connection with developing the crowd gaming framework. Game implementation will not be touched upon further here, since the games are of a relatively simple nature with few truly interesting or innovative implementation details, focus will instead be on the actual framework in itself and how to get from raw input image to a list of glowstick positions from which a mean can be calculated.

For additional details on the implementation specific details of the framework, the complete script called “CrowdFramework” containing all of the programming relevant for the framework part of the project can be found in the Unity project:
APPENDIX F - Unity Crowd Gaming Framework\Unity Crowd Game\Assets\Scripts
On the accompanying DVD.

APPENDIX F also contains all 3 small games as well as their scripts, game assets etc.

6.1 Framework implementation

The framework implementation basically consists of two parts:

- The programming and image processing algorithms responsible for capturing video frames from the modified Playstation Eye camera, color segmenting the frames to separate green and blue glow sticks from the rest of the environment, reducing noise and finally finding the positions of all glow sticks and calculating their mean

- How these final mean calculated positions of the blue and green glow sticks are made available for input in Unity3d authored games

6.1.1 OpenCVSharp image processing

The image processing part of the crowd gaming framework is a step by step algorithm that happens for each frame, the individual steps of the overall procedure is presented here in the order the happen in the framework.
Since openCvSharp is simply a wrapper for OpenCV, it uses many of the predefined and optimized OpenCV functions. Much of the programming presented here is directly transferable to any traditional OpenCv C++ application, with a few changes in syntax.

6.1.1.1 Variable and container declarations

First off a number of variables and image containers are declared for use throughout the script. Only some are included here, and for overview purposes only one of each variable type is included where many variables of the same type are declared:

```
private CvCapture capture;
```

The CvCapture structure is a video capturing structure used only as a parameter for video capturing functions, and does not have any public interface.

```
const int CAPTURE_WIDTH = 640;
const int CAPTURE_HEIGHT = 480;
```

These are simply to integers to hold the desired capture resolution. The const prefix is used since the value should always be the same and is never changed.

```
private IplImage input;
```

A number of IplImages are used throughout the framework, and can basically be regarded as large matrices to hold image data.

```
private CvMemStorage blueBlobs = new CvMemStorage(0);
```

CvMemStorage is a so called growing memory storage. It is a low-level structure used to store dynamically growing data structures such as sequences, contours etc.
It is organized as a list of memory blocks of equal size, the 0 parameter sets the storage size of each of these blocks to a default value of 64 kilobytes of memory.

```
private CvSeq<CvPoint> firstBlueBlob;
```

CvSeq is a growable sequence of user defined elements, in this case CvPoint, which is a 2D set of integer coordinates (x, y).
Many openCv functions, such as FindContours() return CvSeq objects with some elements.

```
private IplConvKernel smallKernel;
smallKernel = new IplConvKernel(3, 3, 2, 2,
ElementShape.Ellipse, null);
```

IplConvKernel are kernels or structuring elements used for morphology operations in OpenCv, by parsing a series of integers as well as an ElementShape parameter, kernels can be customly designed. In this case a 3x3 pixel ellipse shaped kernel with a center in pixel coordinate (2, 2) has been defined.

```
private Vector2 meanBluePosition;
```

Vector2's contain two values x and y, and are conveniently used to store the mean x and you positions of the blue and green glow sticks.

6.1.1.2 Initialization and the Start() function

In Unity3d, the Start() functions is run a single time at the moment the script is activated, in the script a number of initializations that needs to happen at runtime are declared.

```
capture = Cv.CreateCameraCapture(1);
```

First off, the function cvCreateCameraCapture allocates and initializes the previously defined CvCapture structure "capture" for reading a video stream from the camera.
A 1 is parsed to the function as an index for the camera to be used in case the system running the framework has more than one camera available.

```
Cv.SetCaptureProperty(capture,
CaptureProperty.FrameWidth,CAPTURE_WIDTH);

Cv.SetCaptureProperty(capture,
CaptureProperty.FrameHeight, CAPTURE_HEIGHT);
```

The desired width and height property of the frames to be captured are set using the previously define const int variables.

```
input = Cv.QueryFrame(capture);
```

Now the different iplImages are allocated. This happens by first grabbing a frame from the initialized camera and storing it in the iplImage called "input".

```
inputHSV = Cv.CreateImage(Cv.GetSize(input),
BitDepth.U8, 3);
```

To make sure the resolution and formats of all the iplImages match when image data starts being copied back and forth between then, all other iplImages are allocated using functions that return properties of the originally captured image stored in the "input" variable.
The U.8 BitDepth is used making all stored pixel color values be represented by 8 bit integers, allowing for values between 0 and 255. The last number, indicates the number of channel in the iplImage, 3 is used for the input (R, G and B) image as well as the image to hold the HSV

conversion of the image, all image containers handling binary images are given just a single channel of the same bit depth.

```
Cv.NamedWindow("InputStream", WindowMode.AutoSize);
```

Once all the iplImages have been created and allocated, windows to display the images can be created as well. Each window is given a custom name which is used to refer to it and is displayed in the window itself, all windows are set to automatically scale to fit the iplImage it is fed.

```
Cv.CreateTrackbar("Blue MinH", "Blue Thresholded",
blueMinH, 256, ChangeBlueMinH);
```

Finally trackbars to control some of the thresholding variables can be created inside of the windows. The first parameter of the function is a label for the trackbar, second parameter is the desired window in which to display the trackbar, the third parameter is the variable associated with the trackbar, the fourth parameter is the maximum allowed value for the trackbar (the trackbar range will be between 0 and this number), the fifth and final parameter is the function to be called whenever the trackbar is adjusted. Figure 37 shows and example of a window displaying image data and trackbars.

From then on, everything happens once per frame in Unity3ds Update() function.

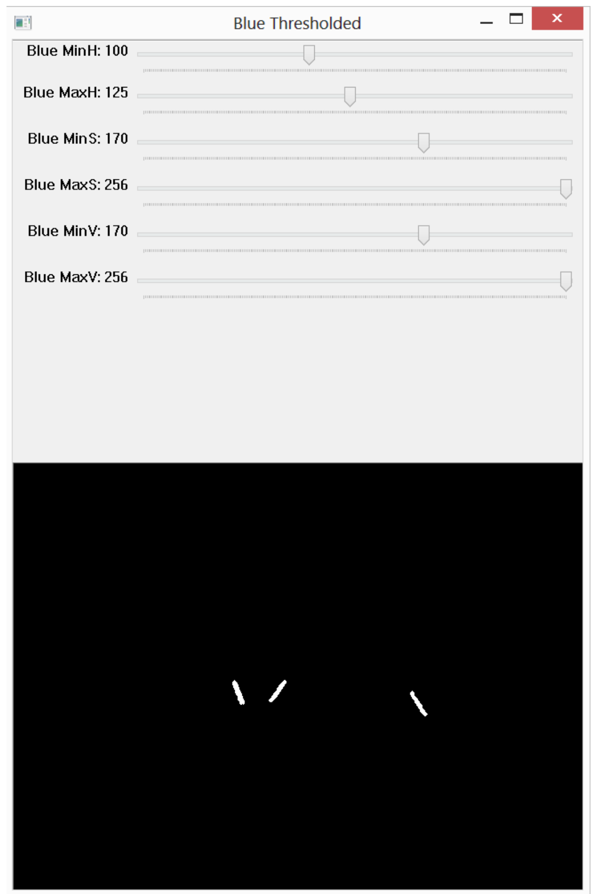


Figure 36: Example of the window displaying the iplImage holding the blue segmented glowsticks. Above the displayed image are 6 trackbars to control threshold values

6.1.1.3 Image acquisition

In Unity, the Update function is called once per frame, it is inside of this function all continuous image processing is handled.

```
input = Cv.QueryFrame(capture);
Cv.CvtColor(input, inputHSV, ColorConversion.BgrToHsv);
```

As was the case in the Start() function, the newest frame is first grabbed from the assigned "capture" instance (camera) and stored in the "input" IplImage.
As investigated in the analysis chapter, color segmentation is preferably done in one of the HS color spaces. Therefore the image is converted from the BGR color space in which it is captured, and into the HSV color space using the standard OpenCv CvtColor function. The original image is kept unmodified and the color converted image stored in another IplImage "inputHSV".

6.1.1.4 Color segmentation

Once the newest frame has been stored and converted, color segmentation using thresholds can be used to separate green and blue glow sticks in the video images from the surrounding environment.

```
Cv.InRangeS(inputHSV,
new CvScalar(greenMinH, greenMinS, greenMinV),
new CvScalar(greenMaxH, greenMaxS, greenMaxV), greenChannel);
```

Here, thresholding of green glow sticks is done using the InRangeS() function also part of the OpenCv library. The function can be compared to a band pass filter, in that it will only let pixel color component values within a certain defined range, pass through from the input to the output image.
The upper and lower threshold values are stored inCvScalar variables, which are containers of up to 4 double variable values, in this case, they are similar to a 3 dimensional vector.
Where the function differs from band pass filters, is in the final output image.
Rather than simply letting a pixels color component value pass through the filter if it is within the defined range, the function will instead evaluate all 3 ranges (H, S and V) and only if all the component values of the current pixel lie within these, the corresponding output pixel is set to white (255) in the binary output image. Figure 37 shows this process.

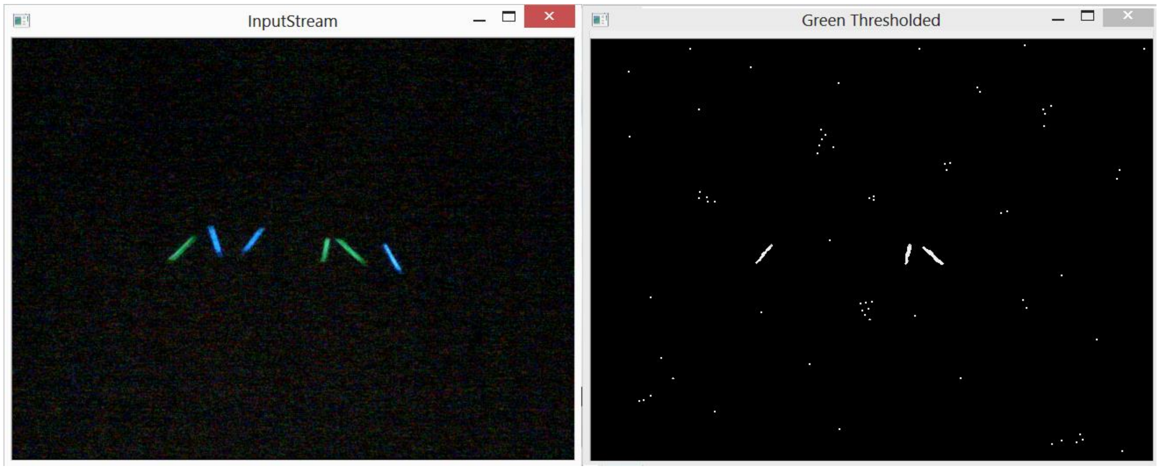


Figure 37: An example of an input image and the resulting thresholded output containing green glow sticks. The image is recorded at very low light conditions similar to those used in the test of the project. This results in a very clear segmentation of the glow sticks (which lights up in the dark), but also some noise from the cameras image sensor.

6.1.1.5 Morphology

Two morphology operations are now run on the thresholded image.

```
Cv.MorphologyEx(greenChannel, greenMorph, tempBlue,
smallKernel, MorphologyOperation.Close, 1);
```

First a closing operation is run to close small gaps in the segmented glow sticks.
Due to the nature of how the glow sticks are constructed, the small broken pieces of glass inside of the glowstick results in it not always emitting light evenly across its entire length. This can in some cases result in small gaps in the segmented objects.
The Closing operation is done using the OpenCv MorphologyEx() function, which takes as a parameters; the source image, the destination image, an image container to temporarily hold image data, a kernel, the desired morphology type and finally the number of times the operation should be performed.
For the closing operation a small kernel (3x3 elliptical) is used, and the operation is run just once and the result stored in a IplImage.

```
Cv.MorphologyEx(greenMorph, greenMorph, tempGreen,
bigKernel, MorphologyOperation.Open, 1);
```

An opening operation is then performed on the same IplImage, to reduce/remove small noisy objects. The operation here uses the same image for both input and output, and uses a slightly bigger kernel. The result of the two operations are seen in Figure 38.

Figure 38: An example of the thresholded image before and after the two morphology operations. Noise is now removed and gaps in glow sticks are patched

6.1.2 Finding glowstick positions and calculating mean

The second part of the framework is that responsible for finding the desired objects in the segmented and cleaned output images, calculating spositions and means, and making these available as input for games.

6.1.2.1 Contour finding (finding the glowstick positions)

For finding the positions of segmented objects, a number of different approaches exists.

1. **Pixel evaluation:**
   If one is simply interested in determining where the majority of objects are located a simple evaluation of all pixels values in the output image may suffice.
   This approach however will have larger objects (glow sticks closer to the camera) weigh heavier on the mean than smaller objects (glow sticks further away from the camera), and since for crowd gaming purposes a number of glow sticks will be present at the same time, this approach is not preferable.

2. **BLOB tracking:**
   A number of algorithms exist for tracking and labeling the individual objects in an image. One approach is the grassfire algorithm, which starts by evaluating the image pixel for pixel. Once it comes across a white pixel it will start a "grass fire" from that pixel, evaluating neighboring pixels, white pixels are "burned" and the grassfire continues until all white pixels connected to the first one are burned. These now constitute what is known as a "BLOB", and a number of further operations can then be applied to the BLOB. These include evaluations of shape (roundness etc.), size (number of pixels), center of mass, bounding box etc. The algorithm can be somewhat slow, and no

standard function for blob tracking exists in the OpenCV library.

3. **Contour finding:**
   What does exist in the OpenCv libraries is a function for contour finding, a useful tool for shape analysis and object detection and recognition.

```
int numberOfGreenContours =
Cv.FindContours(greenMorph, greenBlobs, out
firstGreenBlob, CvContour.SizeOf,
ContourRetrieval.External);
```

   The function retrieves contours from the binary image using the algorithm [Suzuki85] (Suzuki & Abe, 1985) and also returns the number of contours found.

For this project, the contour finding method was used, as this was by far the fastest and most efficient option.
Contours in this case means just the outer edges of the objects, any holes inside the objects are not counted, even though these would also show up in traditional edge detection. The parameters passed to the function are; the source image, a CvMemoryStorage to store found contour data, a pointer to the first element of a CvSeq structure in which the contours will be stored, a header size and the contour retrieval mode (in this case only external contours).

6.1.2.2 Calculating positions and mean

Once all objects have been stored in the CvSeq structures, they can be accessed through the CvSeq interface.

```
for(int i = 0; i <numberOfGreenContours; i++)
   {
      CvRect boundRect = Cv.BoundingRect(firstGreenBlob);
      tempMeanGreenPosition.x += boundRect.Location.X;
      tempMeanGreenPosition.y += boundRect.Location.Y;
      firstGreenBlob = firstGreenBlob.HNext;
      }
```

This loop runs through all of the contours stored in the CvSeq structure “firstGreenBlob”, and add a bounding box around each of them using the OpenCv function BoundingRect().
A bounding box is the smallest box that can fit the entire contour inside of it.
For each new bounding box created, x and y values for its center position are added to a running vector2 counter called “tempMeanGreenPosition”.
Once done with the current contour, the pointer is shifted to the next contour in the sequence using the interface HNext, until none are left.

```
if(numberOfGreenContours > 0){
   meanGreenPosition =
   tempMeanGreenPosition/numberOfGreenContours;
}
```

Finally the running counter to which all bounding box position have been added, is divided by the total number of contours to obtain a mean position. After this the CvMemory blocks are cleared to prepare for the next frame.

6.1.2.3 Showing output and making data publicly available

To be able to see whether or not the image processing functions correctly, the output images of the individual steps are displayed in their respective windows

```
Cv.ShowImage("InputStream", input);
Cv.ShowImage("Green Thresholded", greenChannel);
Cv.ShowImage("Green Morph", greenMorph);
```

The trackbar sliders previously created in the threshold windows, allows for adjusting the framework for each new environment and light conditions.

Outside of the Update function, some custom functions are defined:

```
void ChangeGreenMinH(int newSliderPosition){

      greenMinH = newSliderPosition;
}
```

This function is called whenever the slider for the minimum H value for green glow sticks is adjusted.

Finally the mean positions of glow sticks are made available as input to games using the framework, through a public get function that provides Read-Only access to the data:

```
public Vector2 GetMeanGreenPosition(){

   return meanGreenPosition;

}
```

The function simply returns the latest calculated mean position of green glow sticks as a Vector2 (x, y) coordinate set.

7 Evaluation

In order to evaluate whether the Final Problem Statement has been fulfilled, a test of the designed and implemented games, using the framework, was conducted.
This Chapter contains first a section describing the evaluation in terms of the evaluation criteria that was established.
This is followed by a section on the experiment design and test setup, as well as the reasoning behind these.
Finally the chapter presents the results gathered from the test, and how they relate to each other.

7.1 Evaluation criteria

The Final Problem Statement for this project was as follows:

"How can a Computer Vision based crowd gaming framework, be designed to feature interaction schemes for games, that can support children in collaborative play?"

The Final Problem statement can be evaluated in 3 areas:

- **Computer vision:** was a Computer vision based input framework successfully implemented?

- **Games and interaction:** did the framework perform as intended in relation to its role as input for games?

- **Collaborative play:** were the 3 games designed and implemented in such a way that they were able to facilitate collaborative play when played in the framework?

Each of these areas will be addressed separately in the following sub sections.

7.1.1 Computer Vision evaluation

The performance of the Image processing is of great importance, since it is what provides the input for the games played in the framework. If the implemented image processing and object tracking is not of a robust enough nature it will result in errors or glitches in the interaction with the system. This could be compared to playing a traditional computer game using a broken keyboard or mouse. Intended interactions with the framework may not be recognized, or the framework will respond to interactions where none are. In the end this would result in a confusing and frustrating game experience.

The Image Processing part of the framework can be evaluated on a number of points:

- Does the implementation provide for a sustainable continuous tracking? I.e. are all glow sticks detected consistently from frame to frame?

- How prone is the system to noise from outside coming light sources, colored clothing, projector light etc.?

- How high is the overhead of the framework, i.e. how high was the frame rate of the games when running in the system?

7.1.2 Games and interaction evaluation

As with the Computer vision part of the framework, the way the framework was designed and implemented can be evaluated. This again can be done in a number of areas:

- How does the 3 designed interaction paradigms work in the three designed games?

- Is the simple measure of mean glowstick positions sufficient as input data for interacting with simple crowd games?

- Does the 3 different interaction paradigms relate to the entertainment value, complexity and difficulty of the game experiences?

7.1.3 Crowd gaming and Collaborative play evaluation

Finally the framework can be evaluated in its ability as a crowd gaming framework to facilitate collaborative playing behavior:

- Does the 3 games facilitate collaborative play when played within the framework?

- What group behaviors arise in the playing crowd

7.2 Experiment design

To answer these evaluation criteria, a test was designed around the three games and interaction paradigms implemented.

7.2.1 Test participants

Some days prior to the test, children Sundby Fritidshjem (after school center), were handed notes of parental consent to participate in the test. On the day of the test seven, 3rd grade children (4 girls and 3 boys) aged 8-10 handed back signed notes and these children were accepted as test participants (The signed consents can be found in APPENDIX A on the accompanying DVD). The children were all of normal physical and mental health.

7.2.2 Test setup

The whole test was conducted following a prewritten script which can be found in APPENDIX B on the accompanying DVD.

The test was conducted in a closed off gym hall at the after school center. The hall had a large flat floor, approximately 7 x 5 m., with an elevated platform in one end for a projector to stand on.
A projector and a large screen on the opposite wall, was used to display the game content.
The playstation eye camera required for the framework to function was placed on a tripod in front of the screen, with its cord securely fastened to the floor with tape. Next to that a video camera was placed to record the play sessions.
Participating players were placed in a large group in front of the camera and green and blue glow sticks were evenly distributed amongst them.
Behind the test participants, closest to the projector, a laptop running the framework and games was placed. Figure 39 shows some of the child test participants as seen from the video camera.

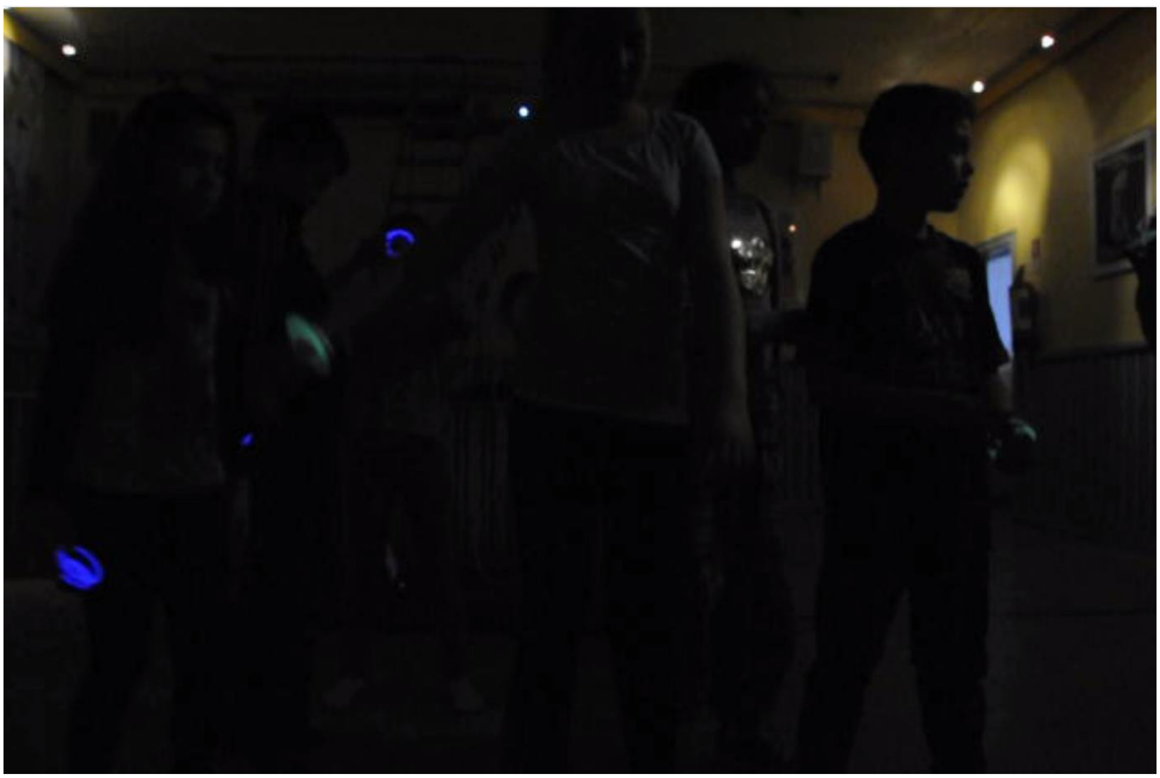


Figure 39: Child test participants as seen from the video camera situated right next to the Playstation Eye camera

A top down illustration of the entire test setup can be seen in Figure 40 on the following page.

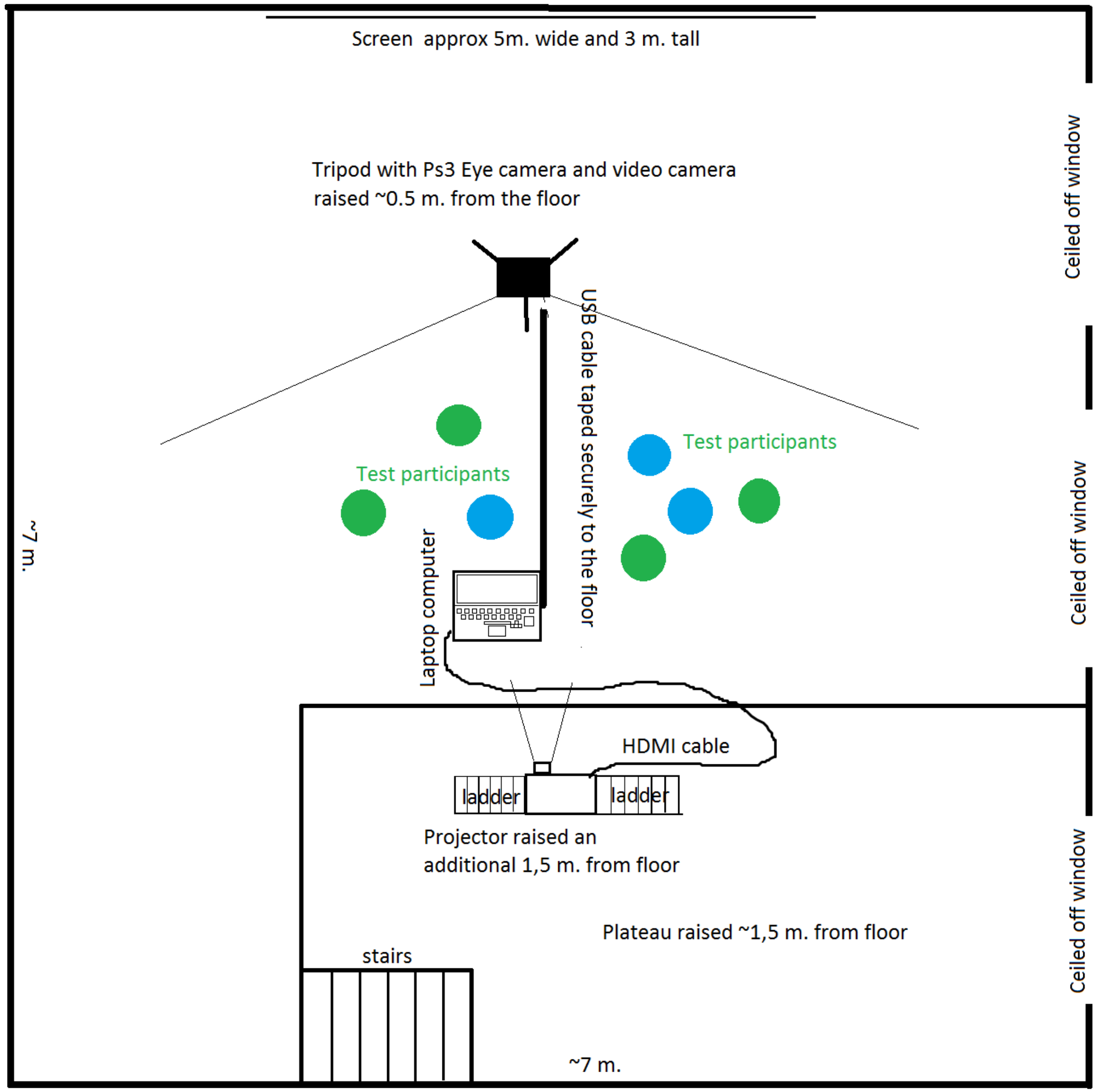


Figure 40: Top down Illustration of the test setup

7.2.3 Test measurements

Three forms of measurements were used during the test. In connection with system monitoring during play, they cover and evaluate the different areas outlined in the Evaluation criteria. Methods and practices for doing usability tests with children were examined in the analysis chapter (see section 4.3, page 54), and some of these where utilized in the design of this test.

The 3 methods used for evaluating the framework were:

- Observation/video recording of the playing crowd and their verbal and physical behavior
- Two 5 point ranking scale type questionnaires for subjective evaluation of the entertainment value and difficulty of the 3 game experiences
- A focus group interview with some of the players immediately following the game experiences

7.2.4 Observation/video recording

Observation of the child test participants as well as video recording of each of the 3 playing sessions is intended to reveal verbal and non-verbal communication in-between the players while interacting with the framework. Movement and gesture behavior can be observed, and verbal communication as well as individual verbal outburst from the players can be interpreted. The General involvement and enthusiasm of the participating players is also readily apparent in this measure, and together with the others add to the overall evaluation of the project.

7.2.5 Scale rankings

Two, 5 point ranking scores for each of the 3 games and interaction designs were designed based on the analysis guidelines for testing with a child target group (See section 4.3, page 54). The scales were printed cutouts, and use smiley icons instead of numbers or text to address two key points of the gaming experience - Entertainment value and difficulty. Examples of the two ranking scales can be seen in Figure 41 and Figure 42.

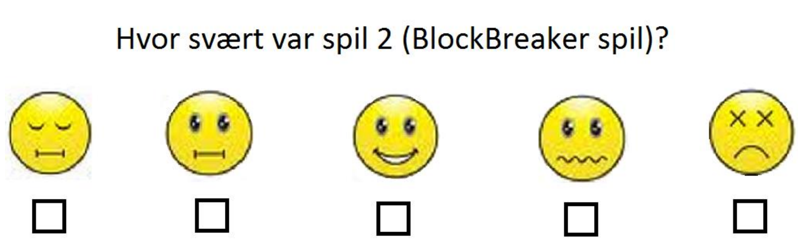


Figure 41: Example of the scale evaluating difficulty of the games. The scales goes from a smiley illustrating "too easy", across "sufficient" and towards "too difficult"

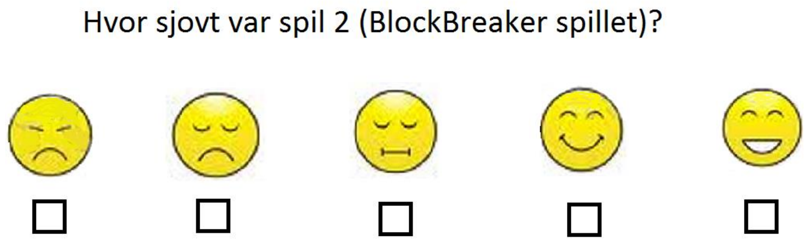


Figure 42: Example of the scale evaluating the entertainment value of a game. The scale goes from a smiley illustrating "not entertaining at all" across "intermediately entertaining" towards very entertaining".

Scales for each game were handed to the children immediately after each game session and instructions to their meaning repeated each time.

7.2.6 Focus group interview

The last measure in the test was a short 10 minute long, loosely structured interview with four volunteering children (two boys and two girls) immediately following the 3 game test session. The interview addresses a number of different topics relating to the game experiences, and was designed and intended to let some of the underlying thoughts and opinions emerge in a casual conversation. The interview was video recorded and can be found (in Danish) in APPENDIX E on the accompanying DVD.

8 Results

This chapter will present the results from the test, in relation to each of the three test measurements. This is done both in terms of their recorded data and how these contribute to answering the evaluation criteria as well as the final problem statement presented (See Section 7.1, page 79).
The section starts with a small note on general system performance based on system monitoring during the test.

8.1 System performance

Based on observations from system monitoring during the test, the framework performed well. The framework and games were running in Windows 8 on a Dell Inspiron 7520 with an Intel Core i7 2,20 Ghz processor and 8 Gb of RAM memory.
With these specifications all games maintained a frame rate of around 30 frames per second (fps) and did not show problems with memory leaks or system freezes or failures.
It is important to note that the more glow sticks that are situated in front of the camera at the same time, the more overhead is required by the framework during runtime.
In spite of their very basic design and implementation, none of the games however had been optimized in terms of performance.

8.2 Observation/Video recording

Due to the fairly dark lighting conditions in the test area and the lack of a camera with night vision or similar light enhancing technology, filming the game sessions proved difficult. Glowsticks however light up in the video recordings, and as such the movements of these are still observable from the video recordings. As is the verbal communication and outbursts from the child test participants. These two factors along with observations by the test facilitator during the test will lay the foundation for these results.
The verbal and physical behaviors are first dissected from the video recordings (APPENDIX D on the accompanying DVD) and divided into 3 tables, one for each game session.
These are then evaluated in relation to the evaluation criteria presented in section 7.1.

8.2.1 Breakout game

Table 2: Observations from the Breakout game session

		Verbal behavior:	Physical behavior:
Relating to general experience:		*“Come on!”* *“Yay!”* *“Noo!”*	Raising arms in triumph Waving/shaking glow sticks vividly
Relating to game rules:		Questions like: *"Are we playing together or against each other?"* *"Why does it (the ball) never come over here?"*	Waiting passively for the ball to pass to their side of the screen
Relating to commands and movement:		*"It’s over here!"* *"This way!"* *"Stay stay (where you are)!"* *"Step back a little"*	Moving entire body left and right Moving glow sticks left and right
Relating to communication between individuals:		*"Elias! It’s your fault, get over here!"* *"Hold it (the glowstick) still!"* *"I can’t help it"* *"Sorry" (when bumping into or stepping on each other)*	Bumping into/stepping on each other (Friendly) grabbing and dragging/pushing each other

Through playing the Breakout game the framework shows great promise as a facilitator of collaborative play. The children participated enthusiastically in the game, and collaborated both verbally and physically to steer their paddles. Participants would shift in and out of leader roles attempting to issue commands on how to move or where to go, and would sometimes physically grab and move other participants.
Collective group or team feelings of success and failure was very apparent in both physical and verbal behavior when hitting and failing to hit the ball.
The children did not seem to have a strong enough notion about left and right to include this into their verbal commands

When it comes to the rules and mechanics of the game, some confusing arose about whether or not the two groups/teams played together or against each other. And some frustration, from especially one of the groups, over sometimes waiting long for the ball to reach a groups/teams side also emerged. This of course was due to the way the game was built with blocks in the middle which when hit will send the ball back to the side from which it just came instead of traversing to the other side like e.g. is the case in *pong.*

8.2.2 Flying game

Table 3: Observations from the Flying game session

		Verbal behavior	**Physical behavior**
Relating to general experience:		Generally enthusiastic outburst like: *“Come on!”* *“This is fun!”* *“Yay!”* *“Noo!”*	Waving/shaking glow sticks vividly
Relating to game rules:		*"What does the yellow ring do?"*	
Relating to commands and movement:		*"Up Up!"* *"Down down!"* *"It’s over here!"* *"This way!"* *"Stay stay (where you are)!"* *"Step back a little"* *"We're right where we're supposed to be"*	Moving entire body left and right Moving glow sticks left and right Holding glow sticks low and high Jumping with glowstick above head
Relating to communication between individuals:		*"Can I say what we do?" (collectively declined)*	

For the flying game may of the same behaviors as were apparent in the Breakout game emerged as well. The game differed in some ways though.
First off, the interaction paradigm has the entire crowd playing together steering the same avatar, but with the groups/colors of players restricted to controlling the avatar on only a single axis. Judging from the verbal outbursts from the children this way of playing was more

entertaining than the Breakout game session. This also has to do with the children being constantly active in this game mode, not having to wait for events in the game to let them play. In terms of group behavior one child actually suggested being "the leader" for this session, telling others what to do, but this was immediately and collectively declined by several of the others.
Interestingly when it came to verbal commands for collective movement the children had a far better notion of "up" and "down" than they had with "left" and "right", and these commands were used actively throughout the game session by the group of children controlling the plane on the y axis.

8.2.3 Racing game

Table 4: Observations from the Racing game session

		Verbal behavior	Physical behavior
Relating to general experience:		*"Its irritating that it keeps turning"* *"What going on here?"* *"No!"* *"This is boring!"*	Turning away from the game in frustration Waving/shaking glow sticks vividly
Relating to game rules:		*"Who (which car) am I?"* *"Why does it keep turning around?"* *"I think it has some kind of a handicap"*	
Relating to commands and movement:		*"All the way up!"* *"There!"* *"Go straight!"*	Moving entire body left and right Moving glow sticks left and right Holding glow sticks low and high Jumping with glowstick above head
Relating to communication between individuals:		*"You have to do like this"* *"Leave me alone"*	

The racing game showed a sudden dramatic shift in both behavior and enthusiasm of the children. Being obviously by far the most difficult of the three games, both player groups had great difficulty making it past even the first turn with their go-karts.

Despite a “dead zone” straight in front of the playstation eye camera, in which the car will only go straight and not turn, both player groups had problems with their car continually turning to one of the sides, making it just spin on the spot. This quickly led to a lot of obvious frustration with the game and the interaction form.

One very apparent problem was that in order to make both cars go straight, all 7 players had to place themselves immediately in front of the playstation camera, or in such an evenly distributed way that the mean of their positions would fall into the previously mentioned dead zone.
This concept, obviously, was much too advanced for the children to grasp during the hectic moments of gameplay, and as such did not work at all.
This also meant that no successful collaborative playing behaviors arose at all during the short time before the game session was terminated. Attempts however to collaboratively steer the individual cars were just as enthusiastic to begin with, as was the case with the two other games. This enthusiasm however, quickly vanished as the game appeared to be unplayable.

8.3 Ranking scales

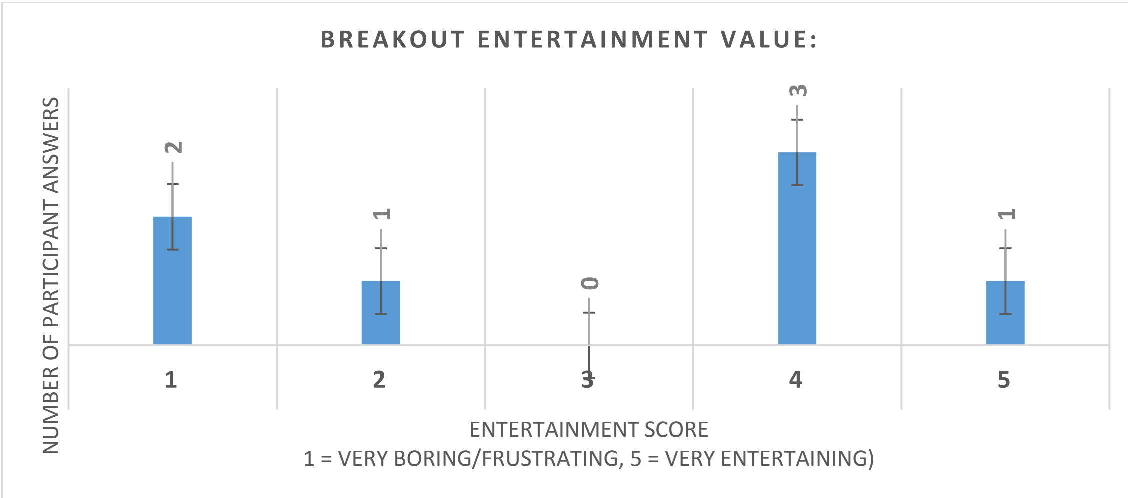
Having dissected the 3 video recordings of the game sessions, taking a look at the ranking scales for each game may possibly confirm or dismiss some of the observations found in those.
The ranking scales are presented for each game in turn.

8.3.1 Breakout game rankings

Ranking scales for the Breakout game yielded the following results divided into a table for Entertainment value and a table for difficulty rating.

8.3.1.1 Breakout game entertainment value

Table 5: Entertainment scores from the Breakout game



What is readily apparent from these scores is the split into two camps, who find the game either entertaining or boring. Upon closer inspection these data relate well to the observations made during the play session with one of the two groups of players spending much play time waiting for the ball to come to their side of the screen. The low entertainment scores does indeed come from these players.

8.3.1.2 Breakout game difficulty

Table 6: Difficulty scores for the Breakout game

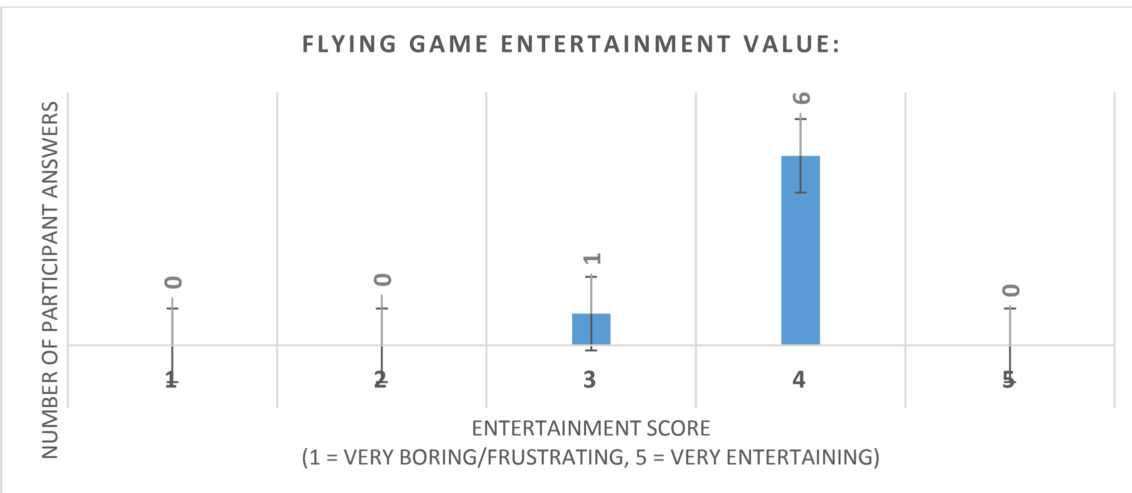
When it comes to the difficulty assessment of the breakout game and its corresponding interaction paradigm, the ranking scales also support the observations well.
The game scores only in the middle region of the ranking scale, corresponding to sufficient/suitable difficulty, with two answers in the mildly difficult and 2 answers in the mildly easy categories. Relating this to the observations made, the children showed little difficulty in collaboratively steering their paddles, and would only occasionally miss a ball. Some of the players in the group consistently receiving the ball may have found this a bit difficult and stressful, while some of the players in the group mostly waiting for the ball, can easily have found this too easy, which would explain the score distribution even more.

8.3.2 Flying game rankings

Ranking scales for the Flying game yielded the following results:

8.3.2.1 Flying game entertainment value

Table 7: Entertainment value of the Flying game

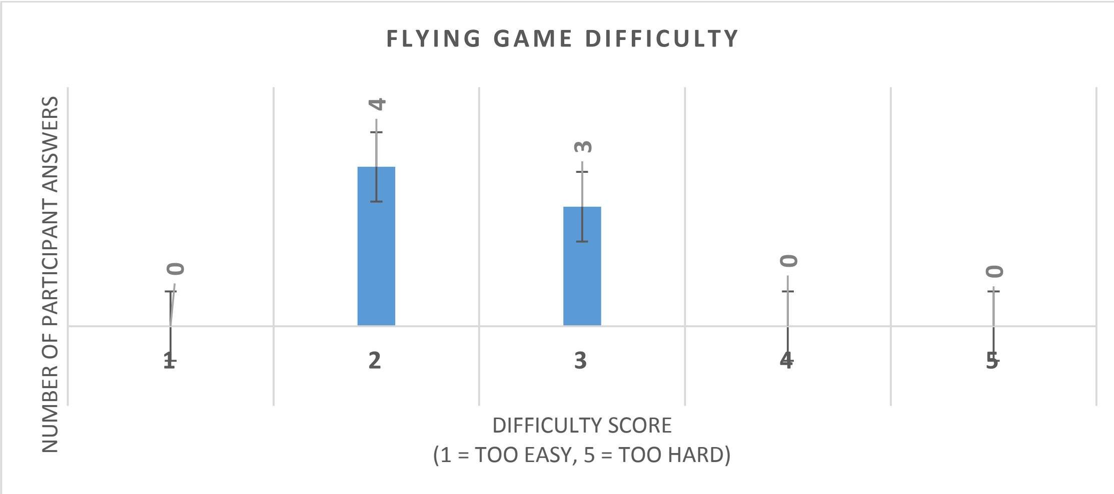


Out of the 3 games, the flying game was the one scoring the highest in terms of entertainment value, with only a single score in the indifferent area and all of the other scores tending towards entertaining.

This corresponds well with the general both verbal and physical enthusiasm observed in the game session. Since this game did not have either group waiting idly for something to happen in the game, the ranking scores does not show the same two-split distribution as with the Breakout game. Instead it shows that practically all the participants found this game and corresponding interaction paradigm to be (the most) entertaining. Moreover, the scale shows that there is no apparent difference in the entertainment value of controlling either of the two axis/directions.

8.3.2.2 Flying game difficulty

Table 8: Difficulty scores for the Flying game



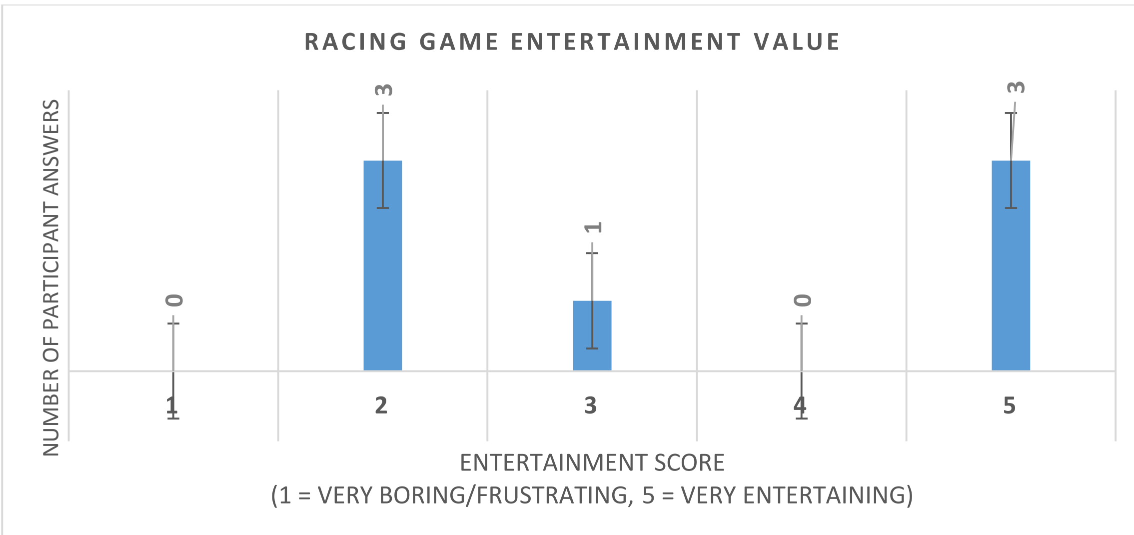
As with the Breakout game, the Flying game was rated as being of an acceptable difficulty, in this case however, slightly more than half of the players found the game to be slightly too easy. Comparing these results with the entertainment scores of the same game, suggests that slightly easy games and interactions does actually result in a more entertaining game experience. Once again the distribution of scores is equal to the group distribution. And upon further investigation, it turns out that players controlling the y axis movements of the airplane, find the game slightly more difficult (though not too difficult) that player controlling the x axis. Most likely this is due to difficulty reaching high enough with the glow sticks, as well as the recognizable nature of the sideways movement interaction by the other group of players.

8.3.3 Racing game rankings

Ranking scales for the Racing game yielded the following results divided into a table for Entertainment value and a table for difficulty rating.

8.3.3.1 Racing game entertainment value

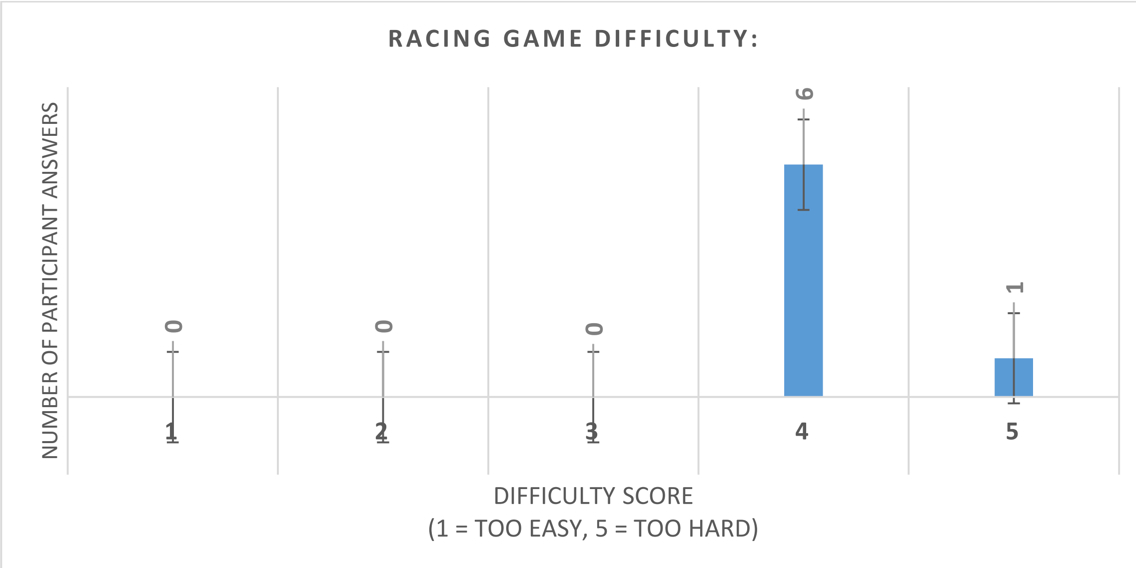
Table 9: Entertainment value of the Racing game



The entertainment scores for the racing game are harder to relate to the observations from the game session itself. It was readily apparent during the game session observations that the game was virtually non-playable by the children in its current implementation and interaction paradigm, and that this led to some frustration from both groups (colors) of players. Despite this 3 of the participants have rated the game experience as being "very entertaining", a fact that will need further investigation in the focus group interview. The other half of the players have, as expected rated the game as being mildly boring, although frustrating immediately sounds more explanatory of the observed game experience.

8.3.3.2 Racing game difficulty

Table 10: Difficulty score of the Racing game



The last of the ranking scores is that of the difficulty of the Racing game, and here the scores are more relatable to what was observed during the game session. All of the participating children have rated the game as being either "hard" or "very hard", as was also what was verbally and physically expressed during the game.
Relating these rankings to the change in interaction paradigm, it becomes apparent that having to focus on moving glow sticks on two axis simultaneously, results in sudden and massive increase in difficulty. The Racing game is the only game of the three that uses this kind of interaction, and coupled with the limited space in front of the camera, this particular interaction paradigm have proven useless.

8.3.4 Focus group interview

To address some of the unanswered questions that arise from the two other test measurements, the loosely structured focus group interview was conducted as a concluding part of the test.
For the focus group interview only 4 of the participating children (2 girls and two boys) took part. A video recording of the interview is found in APPENDIX E on the accompanying DVD.

The interview is in Danish, but to answer some of the questions that arise from the game session observations and ranking scores, relevant questions and comments from the interview are transcribed here in an English translation.

The transcriptions are provided with a prefix "Q:" for Question and "A:" for answer, and are divided into different sections based on what they relate to. Time code for where in the interview the themes (sections) appear are also provided.

8.3.4.1 Why was the car game rated as highly entertaining by some players?

Time code (hours: minutes: seconds): 00h: 00m: 42s – 00h: 01m: 36s

> Q: *"For starters, which of the games did you find the most fun?"*
>
> A*: "The racing game"*
>
> Q: "*The racing game? Ok, why is that? When I was looking the car just kept spinning and spinning"*
>
> A: "*That's what's funny"*
>
> Q: *"Ok, what if you could control it and win over the other team by driving it around the track"*
>
> A: "*That would have been more fun"*

Apparently, what made the game entertaining (and thus score high in entertainment value by some of the participants, was not actual gameplay but simple the comical nature of the car as it kept spinning out of control on screen. All focus group participants agree however that the game would have been more entertaining, had they been able to control the car and e.g. win over the other team by driving it around the track.

8.3.4.2 How does this way of playing/interacting with a game compare to traditional games?

Time code (hours: minutes: seconds): 00h: 01m: 40s – 00h: 03m: 00s

> Q: "*If you compare this way of playing (with the glow sticks) how is it then compared to playing e.g. on your computer or gaming console?"*
>
> A: *"it's like playing on a Wii console"*
>
> Q: "*Like a Wii console alright, have you all tried playing on that?"*
>
> A: "Yeah!"
>
> Q: *"What makes this way of playing resemble a Wii?"*
>
> A: *(Waves arms around to illustrate how to move the glow sticks)*
> *"You* move them from side to side like this, but the Wiimote doesn't light up…
> *It also does not break and spill liquid" (as happened with one glowstick during the test)*
>
> Q: *"How about difficulty? Is it as easy as playing on the Wii?"*
>
> A: *"No it's harder"*

8.3.4.3 Did the system and interaction work as the children wanted it to?

Time code (hours: minutes: seconds): 00h: 03m: 15s – 00h: 04m: 20s

Q: *"Taking as an example the Flying game, were you able to make the plane do as you wanted it to, go up and down and from side to side?"*

A: "*Yeah! – that one was where it worked the best"*

Q: "Ok, what about the tennis (Breakout) game? Were you able to move the paddles?"

A: *"No…, yeah somewhat"*

Q: "*What about in the Racing game? Were you able to steer the car from side to side?"*

A: *"No we couldn't"*

This part of the interview confirms the large differences in usability of the framework over the different interaction paradigms designed. From easily being able to control a single axis, and mostly so in the flying game, where the participants did not so easily clutter in the middle. To the high complexity of having to control two axis simultaneously.

At this point of the interview the participants start to lose focus on the questions and one participant instead inquires to the motivation for creating the games.
After explaining the desire to make many children play together at the same time rather than taking turns, and the potential for using such games in e.g. a cinema while waiting for the move to start, the focus of the interview is naturally led towards themes of how the participants decided collaboratively on what to do.

8.3.4.4 How did the children organize themselves within the groups and decide on what to do when?

Time code (hours: minutes: seconds): 00h: 05m: 20s – 00h: 07m: 00s

Q: *"How did you agree on where go (move your glow sticks)?"*

A: "*We didn't, we just shouted"*

A: *"Michelle just decided everything for me, she grabbed me and dragged me around and waved my arms (waves his arms aggressively in front of him)"*

Q: *"Did you feel most like you were playing individually on your own, or together as a group or a team?"*

A: "*That's a really hard question, we tried to agree on what to do as a whole"*

It is clear from this part of the interview that the concepts of collaborative play are new to the participant, and they are therefore unsure whether or not that was what they were doing. From the observation of the game sessions, a lot of verbal and non-verbal communication took place especially in the first two games, though much of it was unstructured verbal outbursts and physical actions performed by participants very momentarily stepping into a leader role and suggesting/demanding actions from the other players. Judging from the answers here, these were actions of almost unconscious behavior, sparked by the momentarily tense moments in the gameplay, rather than serious attempts of organization within the group.

8.3.4.5 In the eyes of the children, does the framework have potential in the future (would they play again?)

Time code (hours: minutes: seconds) 00h: 07m: 00s – 00h: 07m: 50s

> Q: *"Is this something you would like to try another time, if some of the problems we have talked about were fixed?*
>
> A: *"Yeah, and the glow sticks were fun to play around with"*
>
> Q: *"What about being interrupted while playing and asked to answer questions (the ranking scales)?"*
>
> A: *"That was ok because you get to take a break"*

This final part of the interview addresses some of the potentials for future development of the framework, and whether it has any value at all for playing games. As tokens for interaction with the framework, glow sticks were confirmed as being fun for the children to play with. And reducing some errors as well as removing the all too complex dual axis interaction paradigms, the participating children found the framework to be fun to play with for short periods at a time. This corresponds well with the nature of crowd games as a whole, with small casual games with simple gameplay and fairly short playtimes (between 3 and 5 minutes).
Catching a break in between the intense moments of involvement was a desire expressed through the interview the overall fatigue of the children after the test is clearly apparent towards the end of the interview.

9 Conclusion

The project presented here has been an attempt at creating the foundations of a computer vision based crowd gaming framework for children through which they can engage in collaborative play.
A large part of the project was spend analyzing different concepts in relation to this, eventually leading to a Final Problem Statement.
These include theories of crowd and group behavior, childrens play and the different stages of play, definitions of games and game genres, as well as an extensive State Of The Art review on existing crowd interaction projects.
More analysis of computer vision techniques, crowd activities and how to test with a child target group was also included.

Following the analysis chapters a Design of a Framework involving a modified playstation 3 camera and color segmentation of chemical glow sticks in the video feed provided by this camera. 3 different interaction paradigms were designed around the mean positions of green and blue glow sticks, and 3 games were designed to evaluate the framework through these interaction paradigms.
The framework was implemented in Unity3d using OpenCvSharp for computer vision.

Finally the framework was tested and evaluated through 3 gaming sessions, ranking scale questionnaires and a focus group interview with 7 children from a after school center.

9.1 Concluding on the Final Problem Statement

The Final problem statement was evaluated in relation to its 3 elements:

- Computer Vision
- Games and interaction
- Collaborative play

9.1.1 Computer vision

From a technical Computer vision perspective the Final Problem was fulfilled to great satisfaction. A system was designed and implemented that is capable of recording children holding glow sticks via a modified Playstation Eye camera with a wide Field Of View, The glow sticks can be successfully and continuously segmented from their surrounding environments with very little error or noise. This is achieved through a series of carefully designed steps, many of which rely on robust OpenCv algorithms for implementation.
As a whole the computer vision part of the project works close to flawlessly, with an acceptable overhead and an under a number of different lighting conditions, due to the implemented trackbars that allow the change of threshold values during runtime.

9.1.2 Games and interaction

Based on the segmented glow sticks three interaction paradigms with three corresponding small games were designed and implemented.
A test of the framework with children of an age group corresponding well with the typical age expected for collaborative play to arise, was conducted.
The test showed the great importance of keeping the complexity of interactions through this kind of framework at an absolute minimum.
Through 3 different types of measurements (observations, ranking scales and a focus group interview), the interacting paradigms and games were evaluated both in terms of complexity and the behaviors they provoke, as well as their value as entertainment and difficulty to use.
Two games limiting interaction (glowstick movements), to a single axis was both successful in provoking collaborative group behaviors, and generally scored well in both entertainment value and difficulty rating. The third game, was significantly more complex in its interaction paradigm, which resulted in the game being much too difficult and thus a frustrating and unplayable experience. To answer this part of the Final Problem statement, the interaction possibilities must be kept as simple as possible, while still allowing for collaborative behaviors, and so must the games that make use of them.

9.1.3 Collaborative play

The field of child psychology, and group behavior is vast, and not being a primary field of expertise within the field of Medialogy, it i.e. incredibly hard to encapsulate in a project like this, much more so, to evaluate.
There is no doubt that the simpler interaction paradigms of the first two games presented were successful in provoking several both verbal and physical collaborative behaviors. These were anything from encouragements and shared senses of success and failure, to direct verbal orders and physical interventions between players. In relation to group behavior, examples of attempts at self-organization as well as momentary stepping into leader roles were also observed.
Together with a desire to try to agree on a collaborative effort (expressed in part through the focus group interview), this part of the final problem statement is considered to be somewhat answered. What is important to note is how easily the collaborative playing experience breaks down, when the complexity of the interaction becomes too large. But also how the self-organizing overall movement of a group of children is very hard, due to e.g. the lacking confidence in the difference between the concept of left and right.

9.2 Perspectives

The developed framework provides a solid foundation for expansion in the possibilities for interaction. At its current stage the framework is only capable of segmenting two different colors of glow sticks. Many more colors exist, and the framework could easily be expanded to cover more of these, using other threshold values. For the interaction possibilities (movement of glow sticks) several other interaction paradigms are obvious, without a resulting in an increase in complexity. A simple measure of e.g. the overall activity of waving the glow sticks in the air, is

a simple matter calculating change from frame to frame, either based on individual glow sticks of once again the mean position of all of them. The concept of optical flow, is another area that could be investigated and possibly applied to the framework. Through the measuring of movement direction of the glow sticks from frame to frame, e.g. movement commands or the detection of simple gestures to trigger game events could be implemented.
Finally the inclusion of sound recording and processing is a whole other and interesting field. Since already very much present in the natural gaming sessions already observed, evaluating e.g. the volume or pitch of sounds coming from the players could open for a whole separate set of interaction possibilities. In terms of technical implementations the Playstation eye camera is already equipped with a 4 channel microphone with capabilities for detecting sound direction, while the Unity3d game editor is easily capable of handling simple sound processing algorithms.

10

11 Bibliography

A. I. (1979). Asteroids. *Asteroids*. Atari.

Atari. (1976). Breakout. *Breakout*. Atari Inc.

Atari Incorporated. (1972). Pong. *Pong*.

Avedon, E. M., & Sutton-Smith, B. (1979). *The Study of Games.* Krieger Pub Co.

Barkhuus, L., & Jørgensen, T. (2008). Engaging the Crowd – Studies of Audience-Performer Interaction. *CHI 2008 Proceedings - Work in progress*, (pp. 2925-2930). Florence - Italy.

Berlyne, D. E. (1960). *Conflict, arousal and curiosity.* New York: McGraw Hill.

Birke, A., Schoenau-Fog, H., & Reng, L. (2012). Space Bugz! - A Smartphone-Controlled Crowd Game. *MindTrek*, (pp. 217-219). Tampere, Finland.

Bon, G. L. (2013). *Psychologie des foules.* Puf.

Bregler, C., Castiglia, C., Devincezo, J., Dubois, R. L., Igoe, T., Meyer, J., . . . Wright, B. (2005). Squidball: An Experiment in Large-Scale Motion Capture and Game Design. In *Intelligent Technologies for Interactive Entertainment - First International Conference, INTETAIN* (pp. 23-33). Springer Berlin Heidelberg.

Carpenter, L. C. (1991). *USA Patent No. #5210604, #5365266.*

Chandrasekaran, V., Mohan, S., Pathipaka, O., & Saxena, S. (2004). WeINteract - A Pervasive Audience Participation System. *CHI 2004 - Student Competition*, (pp. 1700 - 1704). Vienna, Austria.

Cinematrix. (2013). *Cinematrix - Interactive Entertainment System*. Retrieved from Cinematrix - Interactive Entertainment System: http://usa.cinematrix.info/

codelaboratories.com. (2013). CL-Eye Test Application. codelaboratories.com.

Farlex. (2013). *The Free Dictionary*. Retrieved from The Free Dictionary: http://www.thefreedictionary.com/crowd

Feldmeier, M., & Paradiso, J. A. (2004). Giveaway Wireless Sensors for Large-Group Interaction. *CHI 2004*, (p. 2 pages). Vienna, Austria.

Foley, J. D., Dam, A. v., Feiner, S. K., & Hughes, J. F. (1990). *Computer Graphics, Principles and Practice.* Addison-Wesley.

Freeman, J. (2008, Volume 7). Large Audience Participation, Technology and Orchestral Performance. *Communications in Computer and Information Science*, pp. 270-283.

Freeman, J. (2013). *Glimmer: Lights, Orchestral Performance*. Retrieved from Leonardo Online: http://www.leonardo.info/lmj/lmj17supp_freeman.html

Fullerton, T. (2008). In T. Fullerton, C. Swain, & S. Hoffman, *Game Design Workshop - A playcentric Approach to Creating Innovative Games* (pp. 42-44). Morgan Kaufman.

Hanna, L., Risden, K., & Alexander, K. J. (1997, September + October). Guidelines for Usability Testing with Children. *Interactions vulume 4 issue 5*.

Hughes, F. P. (2009). Children, Play, and Development,. SAGE.

Madegård, P., & Balslev, J. (2013). *Cinema Dell Arte*. Retrieved from Cinema Dell Arte: http://www.cinemadellarte.dk/

Maynes-Aminzade, D., Pausch, R., & Seitz, S. (2002). Techniques for Interactive Audience Participation. *ICMI '02 Proceedings of the 4th IEEE International Conference on Multimodal Interfaces* (pp. 15-21). Washington, DC: IEEE Computer Society.

McLuhan, M. (1967). *The medium is the message.* Bantam Books.

McPhail, C. (1991). *The myth of the madding crowd.* Aldine de Gruyter.

Moeslund, T. B. (2009). Image and Video Processing. In T. B. Moeslund, *Image and Video Processing* (p. 73). Aalborg University.

Nieboer, G. (2013). *Kids Games*. Retrieved from Gaes Kids Play: http://www.gameskidsplay.net/

OpenCvSharp. (2013). *opencvsharp*. Retrieved from https://code.google.com/p/opencvsharp/.

Parten, M. (1932). Social participation among preschool children. *Journal of Abnormal and Social Psychology 28*, 136-147.

Piaget, J. (1962). *Play, dreams and imitation in childhood.* New York: Norton.

Pick, D. (1995). Freud's Group Psychology and the History of the Crowd. *History Workshop Journal: Volume 40, Issue 1*, pp. 39-62.

Roberts, C. (2011). Control: Software For End-User Interface Programming. *International Computer Music Conference*, (pp. 425-428). Huddersfield, UK.

Roberts, C., & Höllerer, T. (2011, October). Composition for Conductor and Audience: New Uses for Mobile Devices in the Concert Hall. *UIST'11*, (pp. 65-66). Santa Barbara.

Rogers, S. (2010). In S. Rogers, *Level Up - A Guided To Great Video Game Design* (pp. 1-3). Wiley.

Santrock, J. W. (2007). Parten's classic study of play. In *A topical approach to life-span development* (p. 573). New York: McGraw Hill.

Saracho, O. N., & Spodek, B. (2003). *Contemporary perspectives on play in early childhood education.* Information Age Publishing Inc.

Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2009). *Psychology Study Guide.* Worth Publishers.

Scheible, J., & Ojala, T. (2005). MobiLenin combining a multi-track music video, personal mobile phones and a public display into multi-user interactive entertainment. *ACM conference on Multimedia* (pp. 199-208). Singapore: ACM Press.

Schnell, J. (2008). In J. Schnell, *The Art of Game Design - A Book of Lenses* (pp. 24-35). Morgan Kaufman.

Suzuki, S., & Abe, K. (1985). Topological Structural Analysis of Digitized Binary Images by Border Following. *Computer Vision, Graphics and Image Processing*, (pp. 32-46).

ThreeSpeech. (2007). *Playstation Eye - Q+A*. Retrieved from www.threespeech.com: http://threespeech.com/blog/2007/04/playstation-eye-qa/

Unity3d.com. (2013). *Unity3d: Multiplatform*. Retrieved from Unity3d Web site: http://unity3d.com/unity/multiplatform/

Veerasawmy, R., & Iversen, O. S. (2012). BannerBattle - Introducing Crowd Experience to Interaction Design. *NordiCHI 12 - The Nordic Conference on Human-Computer Interaction*, (pp. 228-237). Copenhagen, Denmark.

Verenikina, I., Harris, P., & Lysaght, P. (2003). Child's Play: Computer Games, Theories of Play and Children's Development. *IFIP Working Group 3.5 Conference: Young Children and Learning Technologies.*

VICON. (2013). *VICON*. Retrieved from VICON: www.vicon.com

Vygotsky, L. (1977). Play and its role in the mental development of the child. In J. S. Bruner, A. Jolly, & K. Sylva, *Play: Its Role in Development and Evolution.* New York: Basic Books.

Wright, M. (2005). OpenSound Control: An Enableing Technology For Musical Networking. *Organized Sound*, (p. 17).

12 List of tables and figures

12.1 Figures

12.2 Tables

13 Appendix DVD