

# **Crowd Shape as a Visual Feedback Mechanism in Human-Computer Interaction**

by

Anthony Scavarelli

A thesis submitted to

the Faculty of Graduate and Post Doctoral Affairs

in partial fulfillment of the requirements for the degree of

Master of Applied Science

in

Human Computer Interaction

School of Information Technology

Carleton University

Ottawa, Ontario

© 2015 Anthony Scavarelli

The undersigned hereby recommend to  
The Faculty of Graduate Studies and Research acceptance of the thesis

**Crowd Shape as a Visual Feedback Mechanism in Human-  
Computer Interactions**

Submitted by

**Anthony Scavarelli**

In partial fulfillment of the requirements for the degree of  
Master of Applied Science in Human Computer Interaction

---

Chair, Anthony Whitehead, School of Information Technology

---

Thesis Supervisor, Ali Arya

Faculty of Engineering and Design

School of Information Technology

Carleton University

April 2015

## Abstract

In this thesis, we explore visual interfaces for crowd-computer interactions using a crowd shape generated from participating crowd members, in three main forms: *precise* – a visualization of each tracked individual, *blobby* – an approximate outline or shape of all participants, and *combined* – both the precise shape and the blobby shape layered on top of one another. We wish to determine which of the three crowd shapes is most usable, while also helping determine which kinds of interactions and visual feedback is best suited to crowd-computer interactive installations. We performed usability studies asking participants to rate the *efficiency*, *pleasantness*, *ease-of-use*, and *suitability*, as well as answer an open-ended question about their experience of each crowd shape, for three sample applications. We hypothesized that the combination shape would be the most preferred, with the blobby shape being the next preferred, and the precise shape the least preferred.

We developed 3 applications that tested 3 small crowds of 6-5-6 participants (predominantly 3<sup>rd</sup>-4<sup>th</sup> year university students) for the use of crowd-shape input and visual feedback in collaborative crowd exercises; and 20 (-2 removed) individual participants (predominantly 4<sup>th</sup> year university students and ages 25-34 university graduates) that tested the understanding of crowd-shape from outside the crowd itself. We have found that the blobby shape is the least preferred, and the combined shape likely to be the most preferred. Considering the noise present in the data, we cannot conclusively verify our hypotheses, but the results provide strong indications to the value of both individual and group visual feedback when working with crowds.

## **Acknowledgements**

To my mother who wasn't allowed to finish her education because she was pregnant with me. To her because though the system turned on her she still believed in it and constantly pushed us to explore, engage, and appreciate knowledge and learning as both children and young adults. To her and my father for working so hard to raise their children in an environment that would allow us to succeed even though it was made so difficult for them. Because of their great effort in making sure we had a positive environment in which to grow, I was able to not only finish high school, but also complete a post-secondary and graduate degree. To them I owe great thanks for not only bringing me into this world, but also making me so fortunate as to be loved from the start.

To my beautiful and understanding wife Amy for being so supportive of my exploration in further academic and artistic ventures, and for providing some humor and such wonderful matter-of-fact advice in the moments of uncertainty. For her love. Without her I would be more lost than I already am. In my intrinsic, and undoubtedly frustrating, nature of always questioning my purpose in this universe, I at least know she is beside me. She is my compass, my motivation, and the light unto which I will always aim towards. She is my inspiration and archetype for the person I try to be.

To my children Henri and Eleanor for bearing upon me the importance of unconditional love and being aware of and enjoying the present. The past is deceitful and the future uncertain; but through them I can appreciate the moments of the present in more lucid fidelity – a first step, the first Super Mario level beaten, the first look, the first

puddle. Every moment is the possibility of a new adventure – a new path. Especially for them I try to succeed so that they can be proud of their father, and know that much is possible through adversity – that in fact without adversity the endeavor is lacking. For them I hope to give back the importance in the experience of trying something new. For them I am.

To my supervisor Dr. Ali Arya for impressing onto me the opportunities and exciting challenges of exploring the Human-Computer Interaction field through a lens I was more familiar with – interactive art. For being so supportive of my ideas and my struggles in completing this degree in a timely manner while trying to juggle being an artist, father, and husband. I am grateful for his being a breath of fresh air, for not only understanding of the importance of family, but also impressing its importance actively. Without him I would not be writing this thesis in any present, past, or future moment.

To my friend and artistic partner Henri for being so understanding and supportive while I pursued a graduate degree while we worked together on projects, and for picking up the slack when I was unable.

To the program administrators Erenia and Hana for always having their doors open and to be so quickly in allaying any concerns I may have had through both my undergraduate and graduate degrees. They have been such a huge help in keeping me positive and for always reminding me what is important when asking about my wife and children.

To Dr. Robert Biddle for introducing me to the concept of graduate seminars where the focus is not only on tangible means of evaluation such as assignments and projects

but rather on discussion, understanding, and the search for “wonderfulness” within that understanding. For me this represents the ideals of knowledge gathering and formal education – one that I hope to emulate in any further endeavours as both a student and teacher.

To Dr. Anthony Whitehead for helping indulge my explorations into using affective technologies for meaningful, albeit impractical, person-computer interactions; and of course for his role in creating and shaping the IMD and HCI degrees.

And special thanks to everyone who indirectly and directly helped me get here; including the Social Science and Humanities Research Council of Canada through the IMMERSE Research Network for helping fund my research into gaming ethics and crowd-computing interactions. Thank you to the IEEE GEM and HCI International conferences for allowing me the opportunity to publish and present my work.

And finally, thank you to Carleton University for giving me the opportunity to explore such interesting degree programs, and for the financial support through scholarships and various grants during my several years of school. I started in math-physics and was only able to finish in interactive media and human-computer interaction because they allow second chances. It has been a great experience to be able to research such interesting topics and have others to help discuss and direct their further direction. Pursuing and completing this and my undergraduate degree will be experiences I will cherish indefinitely.

And special thanks to my sister-in-law Dr. Angela Loder for her invaluable and concise feedback in this thesis' final drafts, and for her unfaltering encouragement throughout the entire process of this degree.

# Table of Contents

Abstract.....	i
Acknowledgements.....	ii
Table of Contents.....	vi
List of Abbreviations.....	ix
List of Tables.....	x
List of Figures.....	xi
List of Appendices.....	xii
1 Introduction.....	1
1.1 Research Problem.....	1
1.2 Research Approach.....	3
1.3 Contribution.....	5
1.4 Thesis Organization.....	6
2 Background.....	8
2.1 Reality-Based Interfaces.....	8
2.2 Interactive Installations.....	10
2.3 Crowd Tracking and Visualization.....	12
2.4 Crowd-Computing Interactions.....	13
2.3.1 Cheering Meter.....	13
2.3.2 Crowd Collaborative Classroom Games.....	14
2.3.3 Urban Screen Game.....	15
2.3.4 Lecture Clickers.....	15
2.3.5 Light Around the Edges.....	16
2.3.6. Related Crowd-Computer Interaction Work Conclusion.....	17

3 Crowd Visualization and Application Development .....	20
3.1 Crowd Shapes .....	20
3.2 Test Applications .....	21
3.2.1 Ball-Catch .....	21
3.2.2 Pattern-Match .....	21
3.2.3 Swarm-Chase .....	22
3.3 Application Development .....	22
3.3.1 Software Tools .....	22
3.3.2 Crowd Shape Construction in Experimental Phase 1 .....	25
3.3.3 Crowd Shape Construction in Experimental Phase 2 .....	27
3.3.4 Ball-Catch .....	28
3.3.5 Pattern-Match .....	29
3.3.6 Swarm-Chase .....	30
4 Experiment Design .....	33
4.1 Pilot Studies and Crowd-Computer Development Prototypes .....	33
4.1.1 Early Prototypes .....	34
4.1.2 Final Prototype .....	36
4.2 Experimental Setup .....	36
5 Experimental Results .....	41
5.1 Demographics .....	41
5.2 Quantitative Data .....	41
5.3 Qualitative Data .....	45
6 Discussion .....	48
7 Conclusion .....	52
References .....	54

Appendices.....	60
Appendix I: Demographics.....	60
Appendix 2: Pseudo-Median Point Estimates.....	62
Appendix 3: Questionnaire Response Data .....	64
Appendix 4: Questionnaire Comment Coding Tables .....	66
Appendix 5: Questionnaire for Ball-Catch and Pattern-Match.....	70
Appendix 6: Questionnaire for Swarm-Chase .....	76
Appendix 7: Ethics Clearance Form.....	80
Appendix 8: Email for Call for Participants .....	82
Appendix 9: Consent Form.....	83

## List of Abbreviations

*Box2D*, 2D Physics Engine

*CPU*, Central Processing Unit

*FBO*, Frame Buffer Object

*FPS*, Frames Per Second

*GLSL*, OpenGL Shading Language

*GPU*, Graphics Processing Unit

*IR*, Infrared

*OpenCV*, Open Computer Vision

*OpenGL*, Open Graphics Language

*PC*, Personal Computer

*PS3*, Sony Playstation 3

*P-Value*, Probability Value

*RBI*, Reality-Based Interface

*SDK*, Software Development Kit

## List of Tables

<b>Table 1.</b> List of all hardware technology used. ....	23
<b>Table 2.</b> List of all software tools used. ....	24
<b>Table 3.</b> List of all software technology used. ....	24
<b>Table 4.</b> Mean and Standard Deviation (in brackets) for all Evaluation Criteria.....	42
<b>Table 5.</b> All recorded Friedman P-Values at 0.05 alpha .....	43
<b>Table 6.</b> Point estimates (pseudo medians) .....	43
<b>Table 7.</b> All Wilcoxon signed-rank tests on data pairs for each game type.....	44

## List of Figures

<b>Figure 1.</b> The three shape types used within the Ball-Catch exercise.....	21
<b>Figure 2.</b> The three shape types used within the Pattern-Match exercise .....	22
<b>Figure 3.</b> The three shape types used within the Swarm-Chase exercise .....	22
<b>Figure 4.</b> Steps for creating the blobby shape for Ball-Catch and Pattern-Match. ....	26
<b>Figure 5.</b> Steps for creating the blobby shape for Swarm-Chase.....	28
<b>Figure 6.</b> How the physics bodies are created from the crowd-shape.....	29
<b>Figure 7.</b> Three basic example of how we use image differencing.....	30
<b>Figure 8.</b> Face-trackingin Swarm-Chase.....	31
<b>Figure 9.</b> Prototype Application 1 .....	34
<b>Figure 10.</b> Prototype Application 2.....	35
<b>Figure 11.</b> Physical setups of experiments.....	37
<b>Figure 12.</b> Point estimates (pseudo medians) and confidence intervals .....	44

## List of Appendices

Appendix 1: Demographics .....	60
Appendix 2: Pseudo-Median Point Estimates.....	62
Appendix 3: Questionnaire Response Data .....	64
Appendix 4: Questionnaire Comment Coding Tables .....	66
Appendix 5: Questionnaire for Ball-Catch and Pattern-Match.....	70
Appendix 6: Questionnaire for Swarm-Chase .....	76
Appendix 7: Ethics Clearance Form .....	80
Appendix 8: Email for Call for Participants .....	82
Appendix 9: Consent Form.....	83

# 1 Introduction

## 1.1 Research Problem

Due to the increasing power and flexibility of large projected and screen-based displays, as well as the many types of interaction technologies now easily accessible to even moderately skilled programmers and designers, with frameworks such as OpenFrameworks, Cinder, and Processing [28], there has been a surge of public installation work in recent years within museums, concerts, classrooms, and public outdoor areas. Such installations frequently allow for the interaction of a group of users with a digital system. Interacting with large groups can raise new issues significantly different from those of single user or multi-user cases where each user acts separately. This is related to the notion of gestalt and the unique identity that the whole is independent from its parts - that the individual interactions of each participant in an interactive installation, when combined with the rest of a crowd's interactions create a largely different type of interaction and understanding. [34].

Though touched on in some research, crowd-computer interaction - where a crowd of people whose combined actions interact with an interactive installation or other computing environments - remains elusive. Examples of prior work includes Baarkhuus et al.'s "Cheering Meter" where the culminated volume of a crowds' cheers determined the outcome of rap battles [1], Maynes-Aminzade et al.'s "Pong" where the combined movements of a large classroom move a virtual paddle [8], and O'Hara et al.'s "Red Nose Game" where crowds' movements moved a virtual ball [12]. The main problem

with this type of installation is how we create interaction methods that allow all participants, of which there could be dozens, feel they are contributing to the interaction in one public installation, where feedback is often handled through one large screen. In Chapter 2.4, we will review several methods that researchers have used to address this problem. We generally see two approaches used to create interaction and visual feedback between a crowd of participants and an installation: tracking and displaying every individuals' action using laser pointers [8] or lecture clickers [10], or through computer vision techniques such as head or body tracking [22][23][24]; and secondly determining the crowds' intentions through more holistic methods that combine all individual inputs into a rougher approximation of an action such as the overall crowds audible volume [1] or the crowds general movement [8][12]; which is then displayed as a singular feedback

Tracking all individuals has the advantage that we can then theoretically provide an individual feedback to all participants, so that they all feel directly involved, but it does present technical and usability challenges, particularly as the crowd size increases. Identifying and tracking each individual can be difficult and providing a large amount of individual detail can be distracting and inefficient when dealing with a large group. Combining all individual inputs, on the other hand, has its own advantages and drawbacks. It may be easier to generate collective stimuli but we may lose individual data and involvement. Also, it is not clear what form of visual collective shape is most effective at providing responsive and easily understood input and feedback. We can expect that combining these two approaches can provide benefits in terms effectiveness, pleasantness (satisfaction), ease-of-use (learnability) and suitability (i.e. is the use of the

currently tested crowd shape a valid method of interaction for these particular crowd-computer installations) of the interaction.

In this thesis, we explore the above problems and in particular address the issue of representing a crowd using both collective and individual shapes to provide both input and feedback to the crowd and the individuals that form the crowd. This investigation can help in designing more effective interfaces for systems that interact with a crowd or those that provide crowd information to individuals.

Crowd visualization can also be helpful in automated and semi-automated crowd analysis for various purposes such as security and surveillance, public art, etc. It can be expected that that a more holistic visualization can provide collective information and patterns otherwise not available when looking at individual members. Such pattern recognition tasks themselves are beyond the scope of this thesis.

## **1.2 Research Approach**

In this thesis, we propose a new way of visualizing and identifying a crowd, and an interaction approach that uses “crowd shape” as a variable that can provide specific inputs and also be affected by system outputs. Our research approach is based on design of possible shapes, a series of software prototyping and system development to implement and use those shapes, user studies through various cases, and finally analysis of collected data to compare the performance of possible shapes through a set of evaluation criteria.

We have investigated three methods (*blobby*/approximated, *precise*/people tracked, and a *combination* of these two) of displaying “crowd shape” and studied the usability of these shapes as a means of interaction and feedback. Through these usability experiments, we investigate cases where the user is both interacting as part of the crowd and outside the crowd; with the aim of further illuminating what types of interactions are best suited for crowd-computer interactive installations.

In our experiments, we focus specifically on how the visual feedback and display of a crowd can affect participation effectiveness, pleasantness, ease-of-use, and general suitability as evaluation criteria. We first gather all crowd shape data (individuals’ positions and silhouettes) and combine to create a crowd shape that can be used as input for the system (the three sample applications developed); and then also display the crowd shape (modified) back to the participants, on a screen, as continuous visual feedback to their actions.

The user studies for this thesis are done through three objective-based exercises grouped into two different experiment phases: *Ball-catch and Pattern-Match* for exploring small crowds of 5-6 individuals collaborating together using crowd shapes, and *Swarm-Chase* for exploring how individual participants view and react to a simulated crowd shape of 30 individuals to study their understanding of the crowd shape from an outside or peripheral perspective [2]. These exercises were built to be “continuously variable and socially familiar” [16] with a group-dependent nature to reduce the social embarrassment factor [2], while also utilizing Reality-Based Interaction themes such as “Naïve Physics”, “Social Awareness Skills”, and “Body Awareness Skills” to greatly reduce the gulf of execution (the gap between a user’s goal and the user’s action) for participants [6]. Our

results were based on statistical analysis of quantitative data and also qualitative open-ended responses. Our experimental results suggest that there is a perceived difference between the three different crowd shapes used, and furthermore that the difference is focused on crowd shapes that reflect some feedback as to the individuals positioning in both the Precise and Combined shapes.

### **1.3 Contribution**

The main contributions of this research are to help determine what kinds of interactions are best suited to crowd-computer interactive installations, as well as the type of visual feedback most effective. The following highlight the more precise contributions that stem from our more general research directions:

- Design and categorization of three possible crowd shapes (blobby, precise, combination), in the context of crowd-computer input and feedback methods
- Development of software methods, modifying existing algorithms and processes, for creating the crowd shapes
- Design and development of experimental tasks and performing novel user studies to investigate the performance of crowd shapes based on a set of criteria
- Providing findings and guidelines for design of crowd interaction systems using our proposed Combined method

A paper based on this research has been accepted for presentation in HCII-2015 and publication in Lecture Notes in Computer Science published by Springer.

## **1.4 Thesis Organization**

In Chapter 2, we will overview prior research in crowd-computer interactions, focusing on the general methods that researchers have explored in user interactions and the several techniques used for handling many simultaneous participants. We will also briefly overview interaction principles in smaller scale interactive installations and techniques used in crowd surveillance research. These approaches are useful for gaining information from crowds as both these research fields form a foundational lens unto which we can better understand the current landscape of crowd-computer interactions.

In Chapter 3, we will overview our crowd visualization methods, as well as our process of creating the three applications used for testing.

Chapter 4 is dedicated to overviewing our methods and demographics, including our experimental setup and pilot studies with early prototype applications built, and how these early applications lead to the final applications we used for the experiments.

We will present our experimental results and analysis of our quantitative and qualitative questionnaire data in Chapter 5, and will provide a discussion of the results obtained in Chapter 6.

Finally, some concluding remarks will be provided in Chapter 7. Following these chapters we have all our appendices, including additional information about

demographics, further detail into our results, and the ethics approval and its corresponding materials.

## **2 Background**

Crowd-computer interaction is an amalgamation of many media forms and principles. In this chapter, we will give an overview of related work in the crowd-computer interaction field, as well as Reality-Based Interfaces (the common forms of interaction with public displays) and Interactive installations in general. We also touch upon crowd surveillance techniques that both directly and indirectly have inspired the tracking methods utilized in this thesis.

### **2.1 Reality-Based Interfaces**

The interaction between the participant and digital devices plays a central role in any kind of interactive experience. The usability of an interactive installation, for example, will be defined by not only how well the interactions work for participants but also the interactive method or technology used. As noted in Murphy’s “Lecture Clickers” [10], technology can become a barrier for participation; and as additionally noted by Snibbe et al. “continuous interaction” is a preferred form of interactive design in social immersive media, using full-body interaction so that participants are “unencumbered by electronics or props” [16]. In Jacob et al.’s paper on Reality-Based Interactions [6], the researchers define a Reality-Based Interface (RBI) as one based on interactions that emulate reality with the four themes, which “provides a basis for interaction with computers that is markedly closer to our interaction with the non-digital world” [19] as follows:

- **Naïve Physics:** People have common sense knowledge about the physical world.
- **Body Awareness and Skills:** People have an awareness of their own physical bodies and possess skills for controlling and coordinating their bodies.
- **Environment Awareness and Skills:** People have a sense of their surroundings and possess skills for negotiating, manipulating, and navigating within their environment.
- **Social Awareness and Skills:** People are generally aware of others in their environment and have skills for interacting with them.

Being sensitive to these four principles requires public installation interaction designers to consider basing all interactive modes on forms participants are already intrinsically familiar with due to their similarity to actions they already perform with their bodies in everyday activities. This is because public installations have very little time to impart new knowledge onto participants merely passing by, with assumedly no prior knowledge of the installation. It also encourages installation developers to stay away from technology that may raise the barrier of entry. We see that for large multi-user interactive installations, using full-body control has become a valid and recommended choice based on RBI principles.

## 2.2 Interactive Installations

While research in crowd-computer interaction methods is still in early stages, we can still gain some insight into their likely nature by looking at interactive installations on smaller scales.

The concept of various stages of participant interaction has been revisited in the literature and with varying definitions. Specifically the various stages range from a participant first being in a non-observer role, moving into a peripheral observer role, then to a non-direct interaction role, then finally, and ideally, to a direct interaction role [21][2][17][20][41]. Brignull and Rogers [2] also introduce the concept of social embarrassment, the fear of looking foolish in front of others, being the main barrier to transitioning potential participants from a non-interactive role to an interactive role. Brignull and Rogers mention techniques such as non real-time user input, anonymous interaction, easy-to-use interactive modes, and the “honeypot effect”, the concept of people already interacting attracting others to also try, as potential avenues for reducing the social embarrassment factor [2]. Tang at al. also recommend “provid[ing] low barriers to use, both in terms of fidelity of input and registration of interaction” (similar in objective to RBI principles – see Chapter 2.1); and to have other levels of interaction, such as cell phone or kiosk interaction to “those who were more easily prone to social embarrassment could still contribute ... without making it public.” [17]

Scott Snibbe, an interactive installation designer with over 10 years of experience, and Hayes Raffle, a researcher at MIT Media Lab and Nokia, also lay down an explicit and interesting framework for creating “Social Immersive” installations for use in

museums and potentially other public venues [16]. We, again, see the importance of using reality-based interaction techniques with their reliance on cameras to track body movement; while also defining of several design principles that help guide interesting and effective installations [16]. Specifically these principles are:

- **Visceral:** The media is experienced physically and emotionally, through whole-body interactions, before it is experienced symbolically or rationally.
- **Responsive:** The media responds immediately, clearly and predictably to users' actions.
- **Continuously Variable:** The media, like our natural environment, is continuously changing with infinite variability, usually through simulated response to the user's image, silhouette, sound, location, gestures or other uniquely identifiable features.
- **Socially Scalable:** Interactions are designed to share with others. Furthermore, interaction, representation, and users' engagement and satisfaction should become richer as more people interact. The unbreakable rule is that if the exhibit fits more than one person, it must work with more than one person.
- **Socially Familiar:** The media should augment reinforce existing collocated social behaviors.
- **Socially Balanced:** Interaction equally emphasizes a user's awareness of herself, other users, and the media itself.

Interestingly, there is a common theme within these principles that interactive installations be built with familiarity in mind – very much akin to the core principles of Reality-Based Interactions [6]. We also see the need for interactions to have a very clear and continuous (as opposed to discrete) feedback so that participants are always aware of their effect on the installation

### **2.3 Crowd Tracking and Visualization**

When discussing crowd computer interactions we must also consider the current crowd surveillance models that have been researched and developed to observe and gather information from crowds. More specifically there has been much research into determining whether there are threats within a crowd such as if crowd has entered a panic state [22]. Most of these approaches use computer vision and/or various image processing techniques. These include motion-tracking [24][22] where motion within video is tracked using optical flow, a particle system, using density maps and/or head-tracking [23][24]. More sophisticated methods can also use texture analysis for extracting more holistic properties of the crowd – including “monitoring of traffic flows, detection of disturbances in public spaces, detection of speeding on highways, or estimation of the size of moving crowds”. More holistic methods of analyzing crowds also have the added benefit of maintaining individual privacy by not tracking and displaying individuals [25]. These methods, in the form of points clouds, density maps, or data textures could also make for an interesting addition to any crowd-computer interactions as their analysis techniques could add another dimension of interaction and participant(s) feedback. Examples such as crowd mood or energy state, in future crowd-computer installations for art or

entertainment purposes. It should be noted that the many of the processing and theoretical approaches found in the crowd surveillance literature would be beneficial in building highly responsive real-time crowd-computer interaction systems for very large crowds in the future.

## **2.4 Crowd-Computing Interactions**

In the literature of social psychology and collective behavior, a group or crowd can act as a unique entity. An appropriate example of this comes from Baarkhuus at al.'s "Cheering Meter" crowd-computer installation where it was cited that "as soon as approximately 25 percent of the audience is applauding, the applause quickly cascades to 100 percent" [1]. In this section we will overview the research done thus far with a selection of projects that focus on crowd-computer interaction. To remain identifiably different from the many types of interactive installations, some of which were covered in Chapter 2.2, we will focus on projects where many individuals act as a "crowd" in which the group displays an "illusion of unanimity" [29] as they complete collaborative objectives.

### **2.3.1 Cheering Meter**

Researchers Louise Barkhuus and Tobias Jorgensen [1] created a sound-monitoring system that was manually controlled to receive and measure the amplitude (volume) of the sound generated by a crowd of spectators for rap-battles. This is to determine which of the rap-battle performers received the loudest cheers, and thus arguably the victor of the rap-battle itself [1]. The audience as a crowd is significant as "many crowds are formed as audiences" [3].

By cheering with a large number of others there was no real opportunity for one to see their own or others' individual output; though individuals did express "joy over being part of the concert" [1]. It should be noted that this could also be a weakness in crowd-computer interactions – the difficulty of seeing your own contribution and receiving individual feedback. The cheering meter was lauded by its researchers for its ability to enhance the performance rather than detract from it [1].

### **2.3.2 Crowd Collaborative Classroom Games**

Dan Maynes-Aminzade, Randy Pausch, and Steve Seitz, inspired by a crowd-controlled game at SIGGRAPH in 1991, created: "Audience Movement Tracking", "Beach Ball Shadows", and "Laser Pointer Tracking." [8] Audience Movement Tracking is a game that allowed a crowd to control a paddle's left and right movements in a Pong-type game by leaning left or right in concert. Beach Ball Shadows uses the shadow of a beach ball hit into the air by a crowd cast to deflect missiles from hitting the virtual cities on the ground in a Missile Command type game. Laser Pointing Tracking consists of several games that track many individual laser pointers in the crowd to interact with a projected image. Specific uses of this technology are a "scratching game" that tracks laser pointers to scratch and reveal a hidden image (like a scratch and win lotto card), a graffiti wall that allows multiple coloured lines to be drawn simultaneously, and a whack-a-mole type game that required laser pointers to "catch the moles" [8].

Over eight months, they tested these games on crowds ranging from 150-600 students. Through observations and short surveys completed afterwards by the college-level students, they developed several principles of system design and social factors, such

as “focus on the activity not the game, not required to sense every participant, make the control obvious, play to emotional sensibilities of the crowd, and facilitate collaboration between participants” [8].

Taking in the principles recorded by researchers Maynes-Aminzade, Pausch, and Seitz above, we start to see a crowd-computer interactional framework form, remarking that in crowd-computer interactions it is most important that everyone feels involved even if the technology does not always allow them to be.

### **2.3.3 Urban Screen Game**

Within three UK cities, researchers Kenton O’Hara, Maxine Glancy, and Simon Robertshaw created a camera and projection based collaborative game called “The Red Nose Game.” [12] Each of the three ‘Big BBC Screens’ high above a public space features a camera image of the area directly below it. Superimposed on the camera image are several red blobs. As people walk into the camera image on the projection their bodies are tracked and are able to push around the red blobs into each other so that they can combine, ultimately all combining into one large blob. When all are combined together a point is scored and the game is then restarted [12].

### **2.3.4 Lecture Clickers**

Though Trevor Murphy’s study into using clickers during lectures and for the purposes of greater real-time participation seems very much unlike the many camera-based crowd-computer interactions there are still some significant concepts worth mentioning. This study concerned supplying each student with a clicker during class that

could be used to participate anonymously in to answer questions, give lecture feedback, and collect data during class. Though there were many technical difficulties, the ability to collect this kind of data from the students without revealing who is providing the data allowed greater participation (and less “social embarrassment”) in classroom discussions, particularly ones traditionally controversial such as discussions about religion [10].

Though this paper was much more of an overview of the technology, some interesting lessons do come to surface. In the feedback received from the professors that used the clicker system the drawbacks were described as “difficult coordinating with other faculty to use the same set of clickers, and the clickers could possibly be a distraction to some students” and benefits being “student anonymity, active student engagement, and instant feedback for students on how well or poorly they understand the material.”[10] In the drawbacks mentioned, we can see that having a technology that requires a complex setup, as well as affirming technology that must be dispersed to the “crowd” of up to 180 students can result in issues. Also, as studied and confirmed by Brignull et al., anonymity in public interactions helps keep social embarrassment to a minimal, leading to greater participation otherwise [2].

### **2.3.5 Light Around the Edges**

Todd Winkler’s audio installation “Light Around the Edges” [18] represents a novel way to handle one to many participants. It is included in this survey as his techniques can be scaled up to larger crowds, and brings up some interesting considerations when trying to accommodate crowds of varying and unknown size.

For Light Around the Edges Winkler breaks up his installation into three different modes according to how many participants are present. They are as follows [18]:

**1-4 Participants** - The installation is highly reactive with the speed and location of participants having an immediate and obvious impact.

**5-10 Participants** - The installation's sound becomes that of a train station with players locations triggering related sound effects such as conversations, doors opening and closing, sounds of trains coming and going etc. The interaction is now more subtle.

**Over 10 participants** - The installation's theme becomes that of a big party with participants' movements triggering effects such as canned laughter. The interactions are much more subtle.

Though dealing less with large crowds and mostly focusing on smaller scale interactions that involve only a few participants, some interesting interaction paradigms are defined within this paper, and particularly in the installation "Light Around the Edges". Here the researcher chooses not to create an installation that works the same for all crowd sizes but rather define several modes that are available to be triggered depending on the number of participants interacting.

### **2.3.6. Related Crowd-Computer Interaction Work Conclusion**

Crowd computer interaction is difficult for many reasons; but particularly for its complexity and scale as there are potentially dozens or hundreds of participants to convince are contributing to a particular interaction. In the Laser Pointer and Lecture Clicker studies researchers made sure that all participants could interact and see

themselves as individual entities on a large screen. In the case of classrooms with 600 persons in the Collaborative Games study moving a Pong paddle back and forth, the Cheering Meter, and with the “large modes” of the Light beyond the Edges study, the interaction was based upon the collective input as opposed to the culmination of each of the individuals’ interactions. In this sense we see two main methods of input and feedback in crowd-computer interactions: Tracking and displaying individual input; and tracking and displaying more approximated forms of everyone in the crowds input.

Do all participants feel connected to the interaction when acting as a collective body? And if they do not, is it really that important since the interaction, as in the case of the cheering meter, is merely enhancing their experience of another event (the rap-battle in this case)? Also we feel the use of technology in crowd-computer interactions must be considered. In this sense, we find that most researchers use cameras and image processing software such as computer vision tools to help remove the technology from the hands of the participants and instead rely on more “natural interaction modes” such as voice and the “body awareness skills” [6]. This would make sense as reality-based interfaces do lower the gulf of execution, “the gap between a user’s goals for action and the means to execute those actions” [6]. In this sense we find that there is still no singularly strong method by which to have participants interact together, and feel as if they are participating.

In Maynes-Aminzade et al.’s experiments [8] the researchers espouse the importance of indirect interaction. Specifically the researchers mention not all participants are interacting with the installation itself but rather others, but still feeling involved albeit indirectly - most often through cheers or jeers. In other installations we

are trying to approximate a rough interaction whereby not all have to participate to trigger an action. Both methods leave a bit to be desired as direct interaction for all participants with the installation itself is the ideal mode of interaction. Yet we must also admit that we are looking to trigger a singular action that is a culmination of all inputs from participants. Paradoxically, we want them to feel as if they are participating; but also we need to approximate the input in some form. So we are left with pondering what kinds of interactions are best suited to crowd-computer interactions, as clearly not all modes of interactions are appropriate (e.g. everyone having a keyboard and mouse to in a festival environment). Prior research suggests that reality-based interactions using body movements are more appropriate for large interactions [8][12][16]; and that we should focus on interactions where one large action such as moving a Pong paddle left or right [8] or pushing on a virtual object [12] are the rough approximation and culmination of many individual inputs. Finally how do we handle the visual feedback of a crowd-computer interaction where all can ideally see the reflections of their actions? Succinctly, these are the two main questions we need to explore:

- What kinds of interaction are best suited to crowd visualization?
- How best do we handle visual feedback when dealing with a crowd?

To find an initial answer to the first question, we divided our experiments to the case where a large group of users are interacting with system and the alternative case where a single user needs to interact with a large group. For both question, we investigate the effectiveness of three different visualization methods to explore the concept of using a crowd shape to display input feedback.

## 3 Crowd Visualization and Application Development

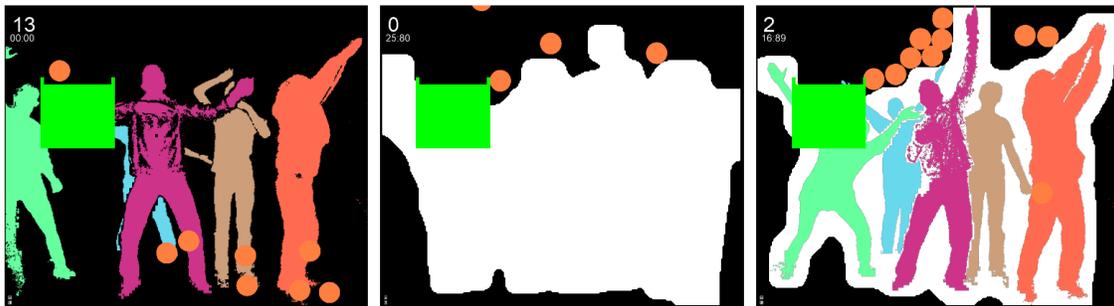
### 3.1 Crowd Shapes

The initial idea of this research was based on the concept of gestalt where a whole or global identity is different from the sum of individual parts [34]. This concept suggests that we should design visualization methods that represent the crowd as one entity without specifying and identifying individuals. After reviewing some related work and some of our own initial tests, we concluded that the role of individual feedback cannot be ignored. As a result, we decided to explore three types of crowd visualization that represent various degrees and arrangements of individuality. Specifically, these are the *Blobby* shape, the *Precise* shape, and the *Combined* shape, shown in Figure 1. While the *Precise* shape represents a typical view of the crowd with emphasis on individuals (shown as silhouettes), the *Blobby* shape aims at visualizing the crowd as a single entity. The *Combined* shape is a layering of both the *Precise* and *Blobby* shape. “*Blobby*” visualization may potentially help viewers understand and interact with crowd as a single element instead of focusing on individual movements. Our primary research hypothesis is that using, or adding such visualization, in addition to a visualization of individuals, creating a *combined* shape, will help improve interaction when individual movements are less important than collective actions. The following sections will describe each of our three experiments that were developed to investigate our proposed crowd shapes. The experiments included three separate and possible scenarios: performing a task using the shape as feedback, performing a task by making a given shape, and finally interacting with a shape created by the computer.

## 3.2 Test Applications

### 3.2.1 Ball-Catch

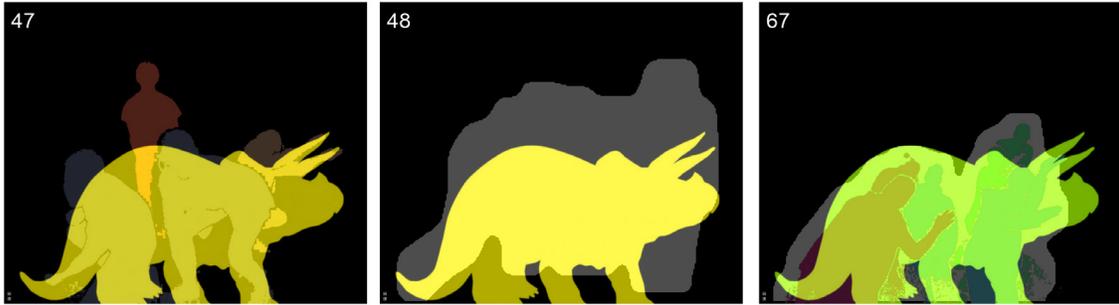
In Ball-Catch, we tasked ourselves with creating an objective-based exercise that would test a group’s ability to form a shape that is most conducive to increasing a score. Specifically the small crowd was expected to work together and form shape that would allow balls falling from the top of the screen to roll into square “basket” (see Figure 1). Once a ball is detected within a basket it is removed from the simulation and the score incremented.



**Figure 1.** The three shape types used within the Ball-Catch exercise. From left to right, Precise, Blobby, and precise/blobby (Combined).

### 3.2.2 Pattern-Match

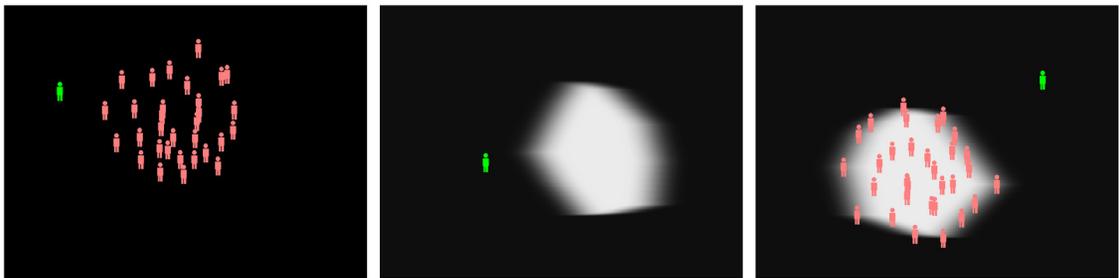
In Pattern-Match we tasked ourselves with creating an objective-based exercise that would test a group’s ability to form to an existing shape. In this experiment we had a score that would represent how closely a crowd of persons’ silhouettes aligned with a shape (a triceratops silhouette) - see Figure 2.



**Figure 2.** The three shape types used within the Pattern-Match exercise. From left to right, Precise, Blobby, and precise/blobby (Combined).

### 3.2.3 Swarm-Chase

In Swarm-Chase we tasked ourselves with creating an objective-based exercise that would test an individual’s ability to understand and interpret a simulated crowd-shape. Specifically, the objective was to avoid a simulated crowd that randomly moves together across the screen. The user and crowd members were represented with small figures on screen, user avatar controlled by the user movements and the crowd controlled by the application.



**Figure 3.** The three shape types used within the Swarm-Chase exercise. From left to right, Precise, Blobby, and precise/blobby (Combined).

## 3.3 Application Development

### 3.3.1 Software Tools

Each program was developed using Cinder Frameworks [28], a C++ and OpenGL coding framework, running on a Windows PC to be displayed on a large 54” TV and

using a Kinect2 for Windows depth camera and SDK [32] for the first two experiments Ball-Catch and Pattern-Match. A Mac PC, a Sony Playstation3 (PS3) Eye high framerate color camera, a large projected screen, and the OpenCV C-based library for live image processing and facial detection [30] were used for the last experiment, Swarm-Chase. Tables 1 through 3 show and briefly describe all hardware and software components used in the creation of this thesis.

**Table 1.** List of all hardware technology used.

Hardware	Description
<b>Ball-Catch</b>	
Windows PC	This was the workstation that was used to develop on and run the application.
Kinect2 Depth Camera	This camera allows us to detect and track individual participants.
60" 1980x1080 Television	This was setup close to participants to visualize the application that simulates a large projected display in a smaller space.
<b>Pattern-Match</b>	
Windows PC	This was the workstation that was used to develop on and run the application.
Kinect2 Depth Camera	This camera allows us to detect and track individual participants.
60" 1980x1080 Television	This was setup close to participants to visualize the application that simulates a large projected display in a smaller space.
<b>Swarm-Chase</b>	
Apple PC	This was the workstation that was used to develop on and run the application.
PS3 Eye Camera	This camera was used to watch and track participants' faces.
1500lumens Projector 1024x768	This was used to project the application onto a large, ~10x8, foot screen.

**Table 2.** List of all software tools used.

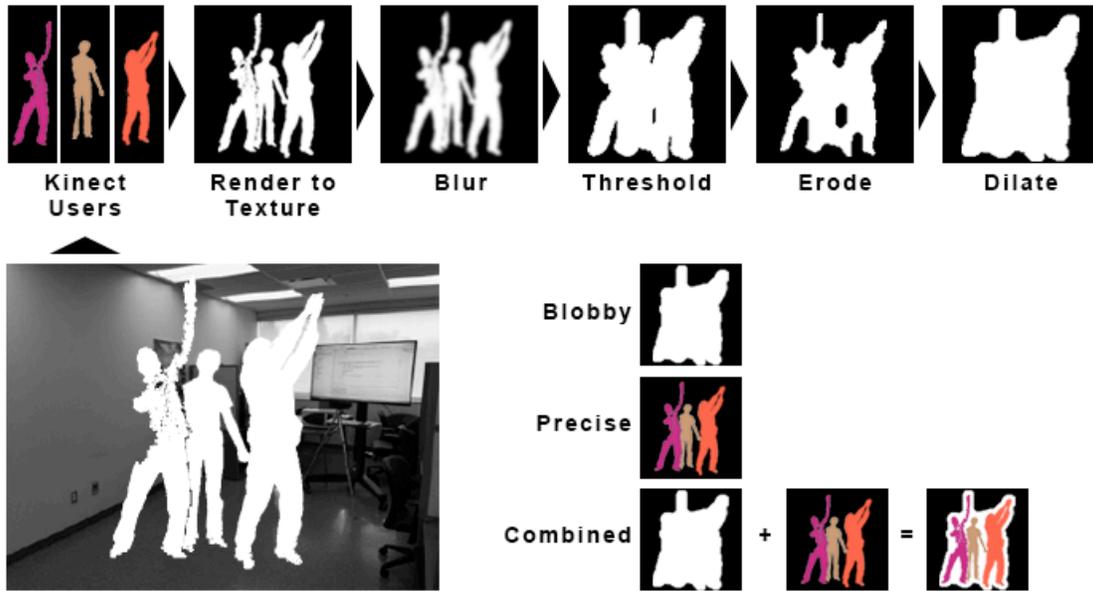
Software	Description
Visual Studio 2013	Programming IDE for Windows.
Xcode 5	Programming IDE for Apple's OSX.
Git / Bitbucket	Source control used for application development.
IBM SPSS	Statistics program used to analyze experimental data.
R / RStudio	Statistics program(s) used to analyze experimental data and create graphs.

**Table 3.** List of all software technology used.

Software	Description
<b>Ball-Catch</b>	
Cinder Frameworks	C++/OpenGL framework targeted for creative programming.
Kinect for Windows SDK 2.0	Library used to interface with all features of the Kinect2.
Delaunay Triangulation	Library created to take a cloud of points and turn into a triangular mesh.
Box2D C++	Library for simulating 2D physics.
<b>Pattern-Match</b>	
Cinder Frameworks	C++/OpenGL framework targeted for creative programming.
Kinect for Windows SDK 2.0	Library used to interface with all features of the Kinect2.
GLSL Shaders	OpenGL's form of accessing and modifying vertex and pixel data.
OpenCV	Library for complex image analysis and object detection.
<b>Swarm-Chase</b>	
Cinder Frameworks	C++/OpenGL framework targeted for creative programming.
OpenCV	Library for complex image analysis and object detection.
Delaunay Triangulation	Library created to take a cloud of points and turn into a triangular mesh.
Flocking	Library created for having objects follow certain rules that give the illusion of autonomy.
PS3 Eye Driver and Library	Library and driver required for using a PS3 camera on a Mac through USB.

### 3.3.2 Crowd Shape Construction in Experimental Phase 1

The first phase of experiments included the test applications where the users acted as a group. The software needed to detect the crowd and then present a visual feedback. We used the Kinect2 depth camera and the Kinect for Windows Software Development Kit (SDK) 2.0 [32] in our Cinder-developed [28] application to easily capture a unique ID and silhouette of each of the 5-6 participants, via Microsoft's underlying skeletal tracking technology. The SDK provides us with programmatic resources for accessing the Kinect2 depth camera and getting tracked user data. After capturing all silhouettes we coloured them a unique colour (see Figure 1, left image), and rendered all participant silhouettes to a single FBO (Frame Buffer Object), an OpenGL render buffer / dynamic texture object that holds both pixel and depth information. We then converted the combined participant silhouettes texture into a single pixel array. From here, we perform blurs, erodes, dilates, and a threshold, using OpenCV image processing functions, to get a blobby shape. This is a very similar process, using similar algorithms, to gathering persistent object tracking data when "blob-tracking" in OpenCV, a process that looks for objects within video frames when using background subtraction techniques. Using the depth camera data instead of colour video frames allowed us to simplify the error-prone process of removing the background from similar foreground objects (for example when a person's t-shirt is the same color as the wall behind them, creating holes in their silhouette when the foreground is removed from the similar background).



**Figure 4.** The steps involved in creating the blobby shape for Ball-Catch and Pattern-Match, using OpenGL and OpenCV.

The main steps we took (as visualized in Figure 4) follow below. For the precise shape we would omit step 4; and for the combined shape we would overlay the precise shape over the blobby shape.

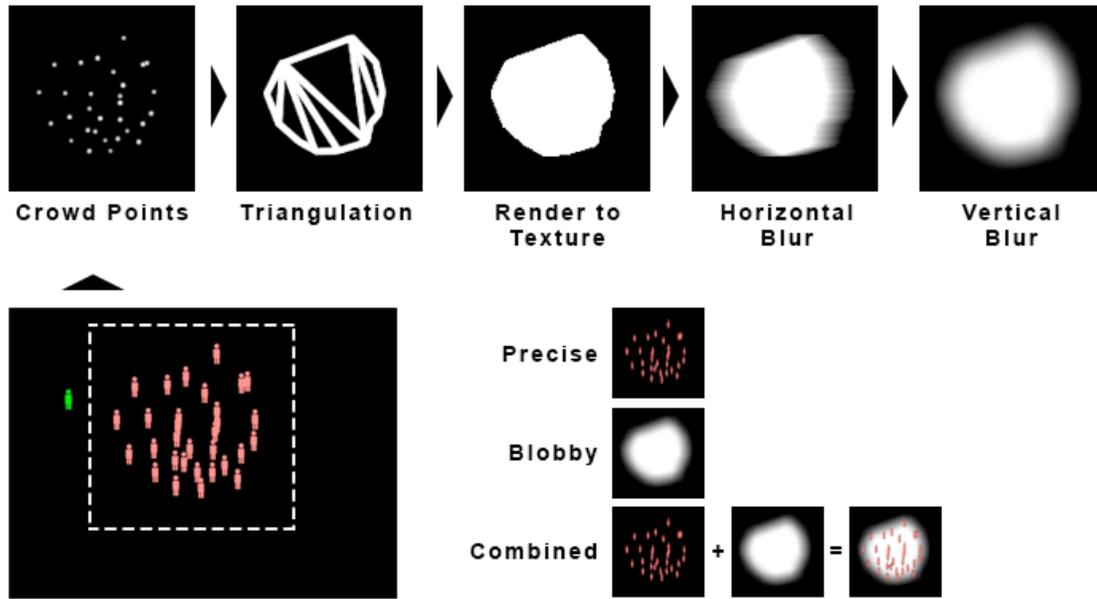
1. All unique user IDs via the Kinect2 depth camera and Kinect for Windows SDK are collected.
2. The body silhouette texture, corresponding to each tracked user ID is then created.
3. We render all user silhouettes into one pixel array.
4. We then passed the pixel array into OpenCV to perform various image processing operations to get a blurry/approximated shape. In particular we converted the image/pixel array to grayscale, blur, apply a threshold (changing image to binary black - no silhouettes - and white image -

silhouettes), erode (reduce the white area of image), dilate (expand the white area of image). We should note that we had tried doing all image processing on the GPU, using shaders, but the speed increase was insignificant due to several instances of ping-ponging (the process of rendering a texture to another texture).

### **3.3.3 Crowd Shape Construction in Experimental Phase 2**

For the experiment phase 2 involving the Swarm-Chase exercise, the process for creating the blobby shape was simplified. In this exercise, participants acted as single individual dealing with a crowd. To create a controlled environment, we simulated the crowd with which the user needed to interact.

The crowd visualization started by creating an array of points representing crowd members. We then created triangles from them, using the Delaunay Triangulation algorithm [31], so that we could then render a plain white polygon to texture, specifically OpenGL FBOs. Using OpenGL Shading Language (GLSL) shaders, for accessing and manipulating pixels on the Graphics Processing Unit (GPU) – a PC hardware component for displaying graphics to a screen, we then blurred the polygonal texture to create more of a “roughness” or approximation, similar to the blobby shape from the Ball-Catch and Pattern-Match exercises. Please see Figure 5 for more detail. This method allowed us to follow similar steps to that found within Chapter 3.3.3 in the previous experiments, even if the initial steps of getting the shape were quite different (simulated versus non-simulated crowd shapes).



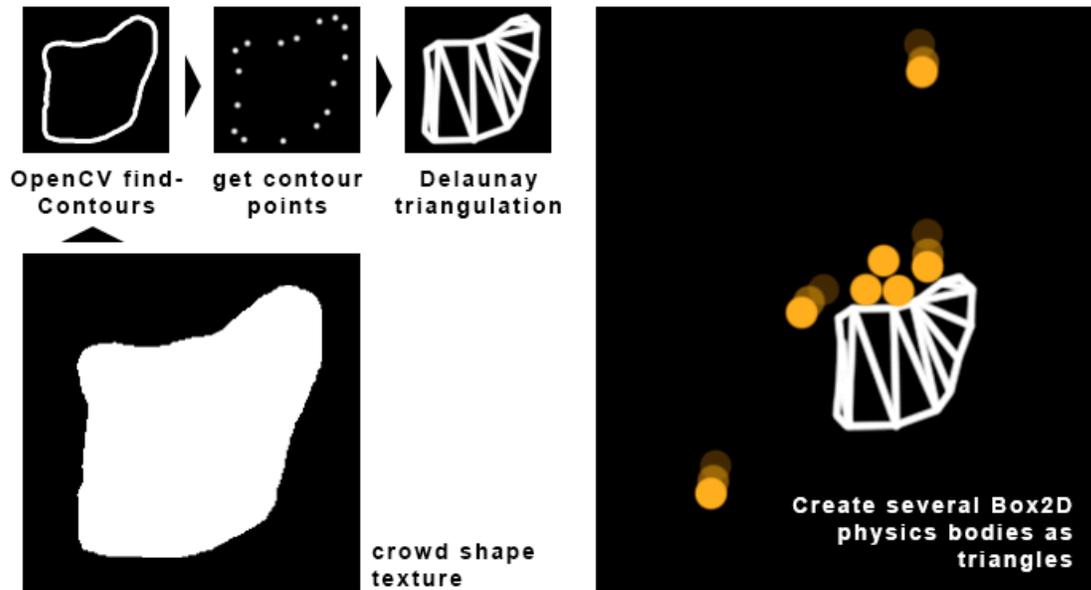
**Figure 5.** The steps involved in creating the blobby shape for Swarm-Chase, using OpenGL, Delaunay Triangulation, and GLSL shaders.

### 3.3.4 Ball-Catch

Depending on which shape we were currently observing (Precise, Blobby, or the Combination) we use Delaunay Triangulation [31] to turn the crowd shape into a series of triangles we could then use as dynamic colliding objects for the falling balls in Box2D. Once a ball is detected within a basket it is removed from the simulation and the score incremented. When using the depth camera we also must make sure there isn't too much direct sunlight as the infrared in sunlight could interfere with the Infra-Red (IR), very low frequencies of light, sensors on the depth camera creating image noise.

When using the Box2D physics engine to create physical objects from the crowd shape that can collide with the balls we have falling from the top of the screen, we must go through a few steps. First we must get the crowd shape texture, and from there we can use the OpenCV findContours() Method to get an outline of the crowd shape. This crowd shape also contains points along the outline we can use to insert into the Delaunay

triangulation algorithms to get several triangles that represent the shape. We can then use the vertices from each triangle to define a custom dynamic body we can input into the physics simulator to collide with the balls we have falling.



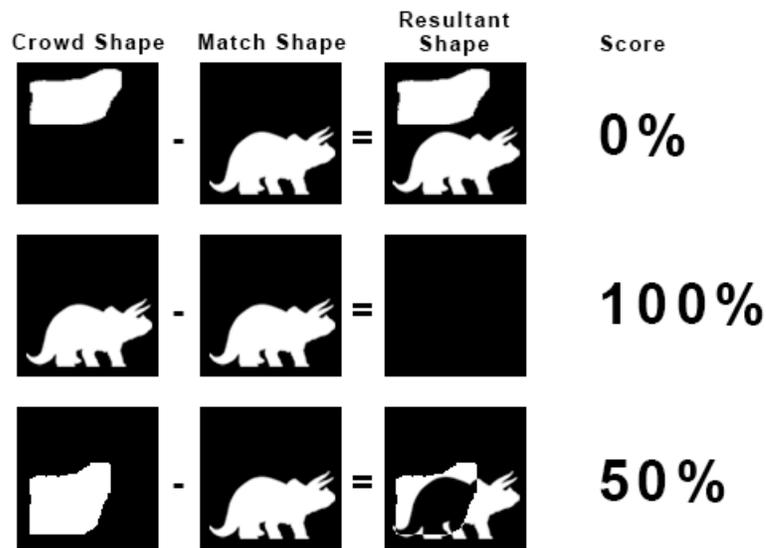
**Figure 6.** Diagram, illustrating how the physics bodies are created from the crowd-shape to allow for balls to collect on crowd shape surfaces.

We use triangles, instead of one large shape, to increase framerate performance as complex shapes can slow down Box2D, and because the Box2D simulation will fail if we try and build concave shapes (i.e. if the crowd shape is a concave shape with “holes”). By using only triangles we are guaranteed to only be inserting convex bodies into the physics simulation (Please see Figure 6 for more information).

### 3.3.5 Pattern-Match

Similarly to Ball-Catch we use a depth camera to capture 5-6 users and their silhouettes as textures we can combine together and process to create a “blobby” shape. To calculate how well (between 0, no match, and 100, perfect match) the crowd shape

matches the triceratops shape we calculate a basic image difference between the two black and white (binary) textures using a GLSL pixel/fragment shader and map the resulting area left to a score between 1-100. Specifically, no difference means a perfect fit, fully black texture, and an fully white triceratops cutout is the least perfect fit possible (please see Figure 7 for some examples).

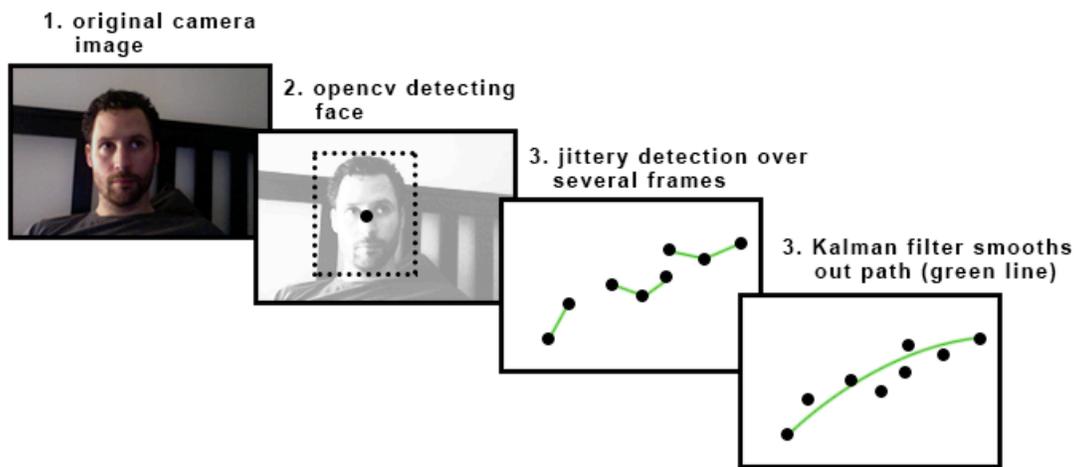


**Figure 7.** Three basic example of how we use image differencing (subtracting one image from another) to calculate a score between 1 and 100 in Pattern-Match.

### 3.3.6 Swarm-Chase

This experiment was greatly simplified in that we only had to track one participant. To further reduce technological complexity we also used a colour camera instead of a depth camera, to track a participants face using OpenCV and a face-detection classifier, smoothing the jittery tracking with a Kalman predictive filter [30][35] (please see Figure 8). This allowed us to build upon and use some of the work developed in the initial prototypes developed in Chapter 4.1.1. Specific advantages of the colour camera over the depth camera include no IR interference or incomplete skeleton construction when

participants are not within camera view, as our OpenCV classifier on the colour images is merely looking for a face and has no need for any other body part within the camera view for reference. We were able to do this as we could assume participants would always be facing the screen allowing us to get a clear image of the face if the camera was positioned at the screen and facing the participant, making tracking far less error prone than that found in our initial prototypes.



**Figure 8.** Using face-tracking and a Kalman filter we are able to get a smooth motion path of a participant in Swarm-Chase.

Once we had a tracked individual, we mapped their movements to a single green person graphic that represented themselves (see Figure 3), and moved in the direction that participants moved. Participant movement was slightly scaled virtually so that they only had to take two steps in in any direction to reach the edges of the large screen that displayed the simulated crowd and individual.

The simulated crowd of 30 individuals was represented by a group of salmon-coloured person graphics; and the crowd's movement was simulated by modified flocking/steering

algorithms [26] that allowed them to move together relatively realistically. We programmed the group to slowly move towards certain targets so that during experiments we could control how and when they move towards the participant's green virtual avatar.

## 4 Experiment Design

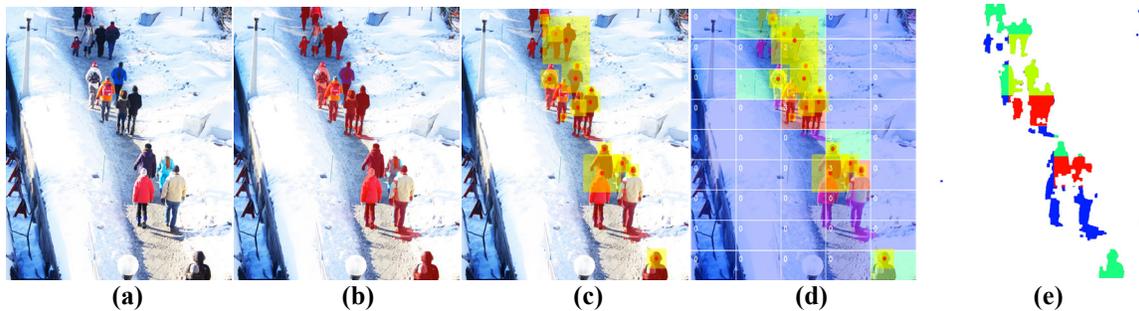
In this chapter, we will overview our initial pilot studies and iterated development prototypes, as well as our experimental setup, demographics, and design for all three exercises *Ball-Catch*, *Pattern-Match*, and *Swarm-Chase*.

### 4.1 Pilot Studies and Crowd-Computer Development Prototypes

When deciding to study crowd-computer interactions our initial thoughts revolved around building a system that could be used for large crowds, for example for use within a classroom or at an event. As we started building the systems, we experienced performance roadblocks that diminished the user experience for our first two prototypes - specifically frame rates unsuitable for real-time interaction and lack of stable tracking of unique individuals over time. Eventually we decided to simplify the setup to study small “crowds” of 5-6 individuals for one phase of experiments (the first two exercises, *Ball-Catch* and *Pattern-Match*) and on individuals within simulated crowds for our second phase (*Swarm-Chase*). By focusing the study on more manageable group sizes we were able to leverage additional technology such as Microsoft’s Kinect2 depth camera for tracking users and less complex classifiers for facial-tracking in our second phase of individual experiments. This allowed us to increase performance to a more acceptable speed that participants found more usable when completing the experiments/exercises. In future studies more research into how to manage and interact with much larger crowd sizes at real-time performance levels would be invaluable.

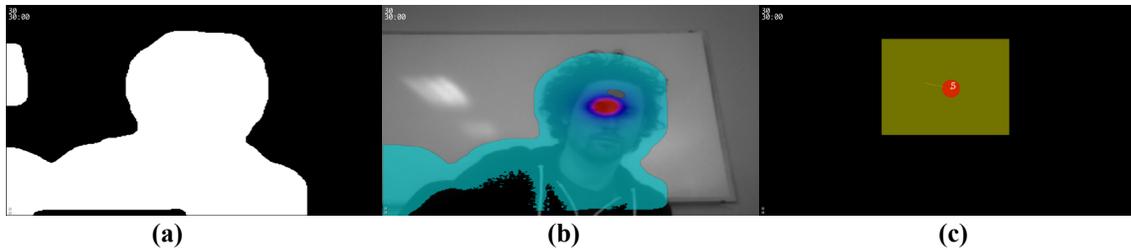
### 4.1.1 Early Prototypes

When setting out to capture crowd data we could use in a real-time interactive application, we initially decided to display the crowd density to better visualize the number and positioning of persons within a relatively large crowd. Our first attempts involved using OpenCV [30], background subtraction techniques, and a trained classifier (a hierarchal data structure created through machine learning processes to look for specific features within an image) to track users' heads. Using the obtained data we would then attempt to overlay a low-resolution density map that would represent the number of heads detected (see Figure 9).



**Figure 9.** In these screenshots of the application (a) we start with an image which is compared to a background image to isolate new people in (b), and then within the resulting foreground shapes a rough head-detection algorithm is applied to detect individuals in (c). From here a grid is displayed of varying resolutions to determine the number of detections per cell in (d) to create a colour/density map where blue represents lower number of detections and red, higher.

Our first attempts at capturing the size and density of crowd proved to be far too slow to be used in real-time situations at an average of 5-10 Frame Per Second (FPS). This was due to complex, and thus slow, head-detection classifier [30] we trained with 3000 unique samples to be able to identify heads in various rotations and positions. Our initial user tests with this setup confirmed our doubts about low performance being a deterrent to the usability of the interaction. Our density map was also deemed not smooth enough to be useful, being unable to track individual participants.



**Figure 10.** In these screenshots of the prototype application (a) we use background subtraction to isolate new “objects” which we assume as people (b), then apply a shader to all detected persons to create a density map and (c) additionally we track persons detected over several frames (the yellow line trail), the size of object (yellow box), current position (red circle), and unique ID (the number ‘5’ in this case).

Our next attempt involved using OpenGL and GLSL shaders to help create a crowd density map that gathered more data from our participants; while also retraining our classifier with less samples so that performance could be increased, at the expense of more false-positives. This tracking system was able to track a detected person over several captured frames (see Figure 10); which also allowed us easily remove false-positives that do not show up for enough frames. This tracking system used predictive algorithms (the Kalman filter [30][35]) and comparisons between detected positions to smooth and better track several individuals simultaneously, by assuming that a track belonging to the same user has the smallest difference in position from the last detected position. Unfortunately this method allowed user tracks to be potentially switched when they overlap each other temporarily. In this attempt, we found that we were able to better track individuals and that the density map was easier to understand for users; but the framerate was merely increased to between 15 and 20 frame-per-second (FPS) which still provided a feedback issue for participants. Finally, we decided to rethink our approach and switch to a depth camera system instead of 2D image processing techniques of colour video images in OpenCV.

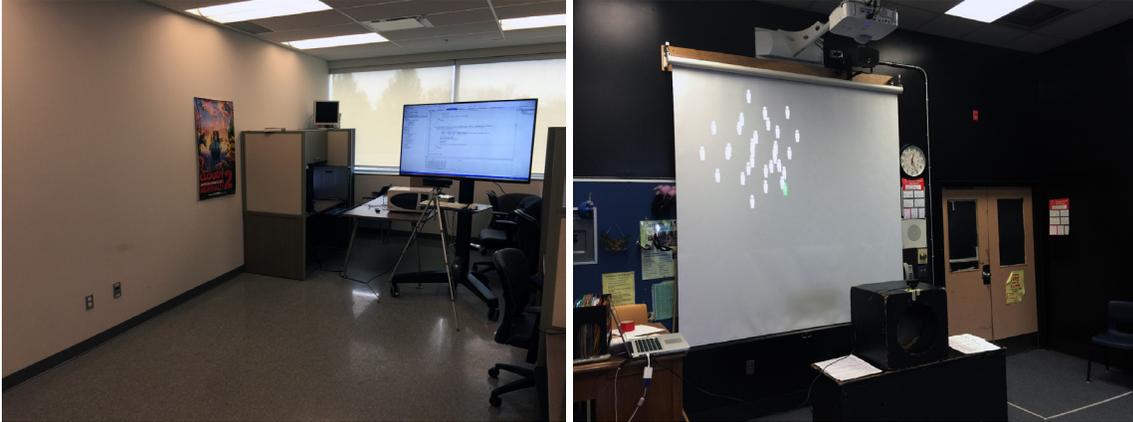
### **4.1.2 Final Prototype**

For our final prototype, which we used for the final experiments, we simplified our experimental conditions to include a smaller crowd size of 5-6 which gave us a bit more freedom to not only focus on the individuals within a crowd but also to use another technology to track them. In this prototype we used a Kinect2 depth camera and the Kinect for Windows SDK 2.0 [32] to track 5-6 unique users and capture their silhouettes and position. This greatly simplified, and enhanced, the silhouette capturing as the depth camera is able to better separate users from the background, as opposed to more traditional background subtraction methods that can introduce errors when users are wearing similar colours to the background being subtracted. We were then able to render the participant silhouettes to textures that allowed us to easily blur, create density maps, and colour for individual feedback (see Chapter 3.3.2 for more detail). This greatly increased performance as the depth camera uses an onboard processor and a specially developed SDK by Microsoft to give user tracking data, instead of using machine-learning to search for persons in a colour image, as used in our previous prototypes using OpenCV [30] (see Chapter 4.1.1 for more details on these prototypes).

## **4.2 Experimental Setup**

The fundamental idea behind this research was that visualization through appropriate crowd shape can improve the users' performance in systems with large number of concurrent users. In order to verify this hypothesis, we started by developing three possible crowd shapes as described in Chapter 3. These shapes represent two possible approaches to visualizing crowds: individual and collective. Our initial pilot tests

showed that users can potentially be interested in both of these approaches so we introduced a third “combined” option that we hypothesized will be associated with the best performance.



**Figure 11.** Physical setups of experiments. The left picture represents the Ball-Catch and Pattern-Match setup and the right the Swarm-Chase setup.

Once the possible visualizations have been determined, we formalized our hypotheses as follows:

1. The most effective, pleasant, easy-to-use, and suitable crowd shape for participants would be the combined shape as it gives both individual feedback and a more general overview of all the crowds' effects.
2. The least effective, pleasant, easy-to-use, and suitable crowd shape for participants would be the precise shape as it gives no clear indication of the crowds' effects.
3. That both hypothesis 1 and 2 would hold true for crowds of participants interacting together (experiment phase 1) and for individuals interacting with a crowd (experiment phase 2).

We then designed two phases of experiments that demonstrate possible forms of crowd interaction: participating as part of the crowd, and participating outside a crowd. First, we explored how several participants work together as a group (in this case 5-6 participants), creating the “illusion of anonymity” [9], to interact with a large screen in two objective-based forms; and secondly an additional experiment that explores how an individual reacts to a crowd shape without the added noise of an actual crowd present – instead focusing on the shape itself, generated from a simulated crowd and crowd movement. Overall we ran three experiments, two for the first 3 groups of 5-6 participants (*Ball-Catch* and *Pattern-Match*) and one for the second group of 20 individual participants (*Swarm-Chase*). To counter-balance learning bias we made sure to randomize which of the two experiments we ran first for each of the three groups in the first experiment phase (including *Ball-Catch* and *Pattern-Match*), as well as in which order we tested each of the three shape types. For the individual experiment phase (including *Swarm-Chase*) we also made sure to randomize the order of the three crowd-shapes we tested. Additionally, *Ball-Catch* was timed (30s to catch as many balls as possible within the basket), as well as *Swarm-Chase* (30s to avoid the roaming simulated crowd); and *Pattern-Match* was only stopped once the group felt they got the highest score they could. We did not analyze the scores recorded by *Ball-Catch* and *Pattern-Match* as 3 scores was not deemed a large enough sample for any insightful analysis; but having the scores did help focus the group’s direction during experiments. *Swarm-Chase* did not track a score.

As public-facing crowd-computer installations are visible to persons of all ages and backgrounds we were not looking for any particular type of participant and chose to

recruit participants using a convenience sampling method where we asked participants, using mass written correspondence and personal approaches to classrooms, studying within the same research building. This included university students within the Interactive Multimedia and Design and Human-Computer Interaction programs at Carleton University. We also received some responses from some high school teachers through recruitment emails to indirect relations when we were unable to find more participants within the school. Please see Chapter 4.3 for more information on our demographics. Financial incentive was in the form of \$5 Tim Horton's (coffee shop) Gift cards.

At the end of each experiment we gave each participant a questionnaire to fill out that asked them to rate on a *7-point Likert scale*, chosen for greater variance, how strongly they felt about their experience across four dimensions: *Effectiveness*, *Pleasantness*, *Ease-of-Use*, and *Suitability*. These dimensions are our evaluation criteria that were chosen for testing user effectiveness, user satisfaction (*pleasantness*), and user learnability (*ease-of-use*) from Neilson's quality components when studying usability [11]. The *suitability* dimension was chosen to gauge participants' opinion on whether using crowd shape as an interactive form is viable for crowd-computer interactions. We also included open-ended questions that asked participants if there was anything they would change about the installation and interaction to gain some insightful qualitative data. The results of both the quantitative and qualitative analysis can be found in the results (Chapter 5); and the entire questionnaires themselves found in Appendices 4 and 5.

To analyze the data, we compare each individuals' survey responses, for both the group-based experiments and individual experiments, for each of the three crowd shapes, and use the Friedman statistical test, as we assume our data is ordinal and we are comparing three(greater than two) related samples (each of the four measures we are testing from each crowd shape, from each participant). After determining if there is significance (less than or equal to a 0.05 p-value) within our 3 sample tests we then used the pair-wise Wilcoxon ranked-sum test because they are related samples, to find specifically where that significance lies between the 3 crowd shapes. To analyze our qualitative data from the questionnaires we performed a simple coding by determining four most prevalent (most common frequency) themes within each exercise and note what was specifically concerning the observed theme. From these notes we then try to infer why these four themes were the most prevalent.

## **5 Experimental Results**

### **5.1 Demographics**

The first phase of experiments included 3 groups of 6-5-6 members, giving us 17 participants in total for the “group” experiments. In the “individual” experiment, Swarm-Chase, we had 20 participants. For both phases we had approximately 2/3 males and 1/3 females, and most participants were in the process of, or had completed, a post-secondary degree. Ages for the first phase were almost completely within the range of 18-24 (predominantly university students) while the second phase expanded to include most of its participants within the ages of 18-34 (predominantly university students and young secondary school teachers), with 22% ages 35-54 and another 22% 55+. Computer expertise was slightly higher in the lower age groups. It should be noted that in the Swarm-Chase two participants’ data were removed as the questionnaires were filled incorrectly and their comments afterwards about “hoping they answered how the researchers would like” suggesting a possible Hawthorne effect where they were not answering for themselves [18]. This brought down our sample total for Swarm-Chase from 20 to 18. Please see Appendix 1 for more detail on our collected demographic information.

### **5.2 Quantitative Data**

Table 4 shows the mean and standard deviation for responses on our evaluation criteria in three experiments.

**Table 4.** Mean and Standard Deviation (in brackets) for all Evaluation Criteria.

	Effectiveness	Pleasantness	Ease-of-Use	Suitability
<b>Ball-Catch (n=17)</b>				
Blobby	4.47 (0.80)	4.71 (1.49)	4.29 (1.36)	4.29 (1.31)
Precise	5.12 (1.45)	5.59 (1.37)	5.12 (1.45)	5.00 (1.66)
Combined	4.94 (1.09)	5.29 (1.36)	5.12 (1.17)	5.41 (1.23)
<b>Pattern-Match (n=17)</b>				
Blobby	4.35 (1.50)	4.24 (1.60)	4.18 (1.55)	3.71 (2.02)
Precise	5.41 (0.94)	5.82 (1.07)	5.29 (1.65)	5.47 (1.59)
Combined	5.12 (1.11)	5.53 (1.07)	5.29 (1.21)	4.88 (1.36)
<b>Swarm-Chase (n=18)</b>				
Blobby	5.11 (1.18)	5.06 (1.55)	5.56 (1.38)	5.33 (1.24)
Precise	5.61 (1.14)	5.72 (0.89)	5.89 (0.96)	5.44 (1.25)
Combined	6.00 (1.08)	5.78 (1.06)	6.11 (0.96)	5.72 (1.27)

As these results were obtained from Likert scale data (ranges 1-7) we can only assume that the data is both ordinal and non-normal leading us to analyze the data using non-parametric statistical methods. Since we used repeated testing procedures with the same participants to collect the data respective to each of the three crowd shapes (*Blobby*, *Precise*, and *Combined*) across four dimensions (*Effectiveness*, *Pleasantness*, *Ease-of-Use*, and *Suitability*), we used a Asymptotic Sig. (2-tailed) Friedman test to determine if there is significant variance between the crowd-shape response data. The results are shown in Table 5.

After determining where there are differences between the three dependent groups of data we conducted further post-hoc testing using the Wilcoxon signed-rank test to

determine where the significant differences lie between each pair (*Blobby-Precise*, *Blobby-Combined*, and/or *Precise-Combined*). Please see Table 5 for each Friedman p-value calculated at alpha 0.05; and Table 7 for p-values obtained between each pair using the Wilcoxon ranked-sum test.

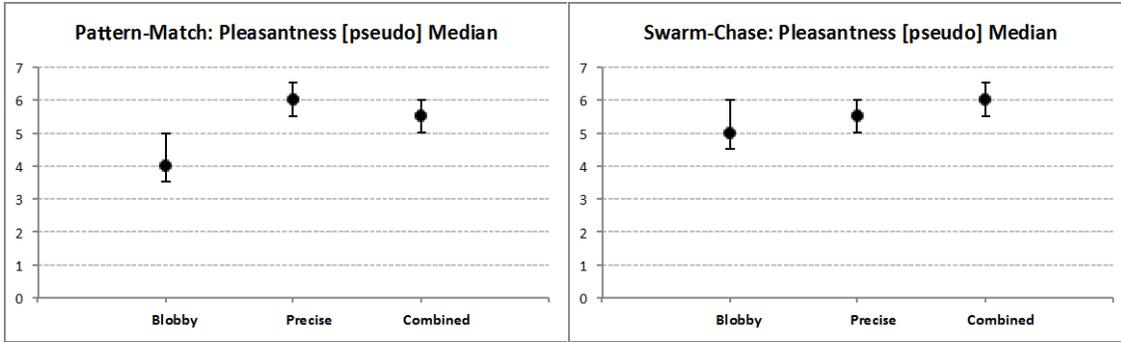
In Figure 7, we can see two examples of the point estimates / pseudo-medians of both Pattern-Match and Swarm-Chases' graphed with their appropriate confidence intervals calculated in R.

**Table 5.** All recorded Friedman P-Values at 0.05 alpha - Asymptotic Sig. (2-tailed) with degrees of freedom of 2. Those highlighted cells represent where we see significant differences between the response data i.e. we reject the null hypothesis that the samples are the same.

	Effectiveness	Pleasantness	Ease-of-Use	Suitability
Ball-Catch (n=17)	0.118	0.309	0.576	0.341
Pattern-Match (n=17)	0.021	0.001	0.017	0.028
Swarm-Chase (n=18)	0.003	0.027	0.05	0.285

**Table 6.** This table presents all the point estimates (pseudo medians), consistent with the Wilcoxon test. We ignore Ball-Catch as there were no significant differences detected by the Friedman test.

	Effectiveness	Pleasantness	Ease-of-Use	Suitability
<b>Pattern-Match (n=17)</b>				
Blobby	4.5	4.0	4.5	4.0
Precise	5.5	6.0	5.5	5.5
Combined	5.0	5.5	5.5	5.0
<b>Swarm-Chase (n=18)</b>				
Blobby	5.0	5.0	5.5	5.5
Precise	5.5	5.5	6.0	6.0
Combined	6.0	6.0	6.0	6.0



**Figure 12.** Visualized here are the point estimates (pseudo medians) and confidence intervals consistent with the Wilcoxon signed-rank test that show possible differences between the Bloby, Precise, and Combined shapes for the pleasantness response for the Pattern-Match (left) and Swarm-Chase (right) experiments.

Looking at the point estimates in Figure 12 we can see that there should be some differences between responses and so Wilcoxon signed rank tests were performed on all sets that passed the Friedman test, using a Bonferroni adjustment of the p-value from 0.05 to 0.017 (0.05/3 samples) to determine significance in Table 7.

**Table 7.** All Wilcoxon signed-rank tests on data pairs for each game type (except Ball-Catch as no Friedman significance detected). If significant i.e.  $p < 0.017$  (via Bonferroni adjustment to reduce a type 1 error) then the cell will be highlighted.

	Effectiveness	Pleasantness	Ease-of-Use	Suitability
<b>Pattern-Match (n=17)</b>				
Blobby-Precise	0.042	0.003	0.03	0.014
Blobby-Combined	0.028	0.001	0.004	0.012
Precise-Combined	0.34	0.26	1	0.142
<b>Swarm-Chase (n=18)</b>				
Blobby-Precise	0.075	0.085	0.222	0.642
Blobby-Combined	0.002	0.015	0.026	0.107
Precise-Combined	0.052	0.862	0.234	0.132

Looking at our results we see that there are some perceived differences and patterns within the data. We can see that the Friedman test exposes some sample fluctuations in

Pattern-Match and Swarm-Chase across all four dimensions, with the exception of suitability in Swarm-Chase.

Within Pattern-Match we can see that after the Friedman tests expose differences within the samples, that using Wilcoxon post-hoc tests, show the differences tend to lie between the Blobby and Combined shapes for both Pleasantness, Ease-of-Use, and Suitability. Interestingly the Combined shape is considered more pleasant, more easy-to-use, and more suitable than the Blobby shape in Pattern-Match. We can also see that the Precise shape is rated higher than Blobby in both Pleasantness and Suitability. We can see this result echoed in Swarm-Chase where the significant differences within the sample lie between the Blobby and Combined shapes in both Effectiveness and Pleasantness. In Swarm-Chase the Combined shape is deemed both more effective and pleasant.

### **5.3 Qualitative Data**

Overall we observed that participants seem to enjoy themselves during the crowd-based games: Ball-catch and Pattern-Match. Ball-Catch did seem to incite more critiques about the technology in the comments where we found 4 unique mentions of lagging or low frame rate; and during play the “slowness” was often attributed as an issue to enjoyment and objective particularly when using the Combined shape.

Also, participants seem to generally prefer a shape that had the Precise or “people” shapes present, as there were positive reflections (comments that use wording such as “prefer”, “liked”, or “best”) within the comment areas about both the Precise and Combined shapes; though interestingly none about the Blobby shape exclusively.

Additionally, when looking through the data we do see references to participants enjoying the combined aspect, but critiquing its size, with comments such as it “was easier to get the approximate shape but the blob outline was too large”, “method was pretty good but maybe make the Blobby shape a bit smaller”, “Best. Would be nice to have a less wide shape”, “Overall I found it best”, and interestingly “favourite b/c I could see the ‘people’ and the boundaries ...”.

Participants also came across as quite interested in the interaction as many suggested possible changes to the shapes and to the game itself with 26 mentions of changes to make the game(s) better and 42 mentions of how to make the shapes better within the comments area, focusing on the Blobby shape in particular as a source of ambiguous visual design: In the comments we see sentiments of this where the Blobby shape was mentioned as showing “little information on what the blob represents”, “not resembling much different human beings”, needed “more stimuli to represent the crowd”, and “hard to distinguish who is where”. Fortunately, only 3 comments were made about how either the shape or exercise was confusing over all experiments.

If we look to Appendix 4 to the coded qualitative data we can see that the four main themes observed in the Ball-Catch were a preference for the combined shape (6 mentions), technical Issues (4 mentions), critique of the blobby shape itself(14 mentions), and critique of the game mechanics themselves(10 mentions). For Pattern-Match the most common themes observed were a preference for the combined shape(6 mentions), technical issues(6 mentions), the blobby shape itself not viable(4 mentions), and the shape to match (the triceratops) being too complex (3 mentions). For Swarm-Chase we see the four main themes being observed as a preference for the combined shape(3

mentions), technical issues (6 mentions), critique of the blur on the blobby shape(6 mentions), and critique of the game mechanics themselves (6 mentions). No discernible differences were found between genders, education levels, or computer expertise.

## 6 Discussion

Looking at the results, we can see that the Blobby Shape seems less effective, less pleasant, less easy-to-use and less suitable than the Combined shape in all areas we determine there is significance in Pattern-Match and Swarm-Chase. Additionally, we can see that the Precise shape is also determined to be more effective, pleasant, easy-to-use, and suitable than the Blobby shape as well. There is no significant difference detected between the Precise and Combined shapes, although the comments received also seem to suggest some preference towards the Combined shape. Concerning our three hypothesis from Chapter 4.2, we see that the combined shape does indeed seem to be suggested as the most preferred shape amongst participants, when referring to the open-ended question data; but statistically we can only say that either the Precise or Combined shape is the most preferred across our four measured dimensions for both Pattern-Match and Swarm-Chase. Future studies would likely best focus here with larger sample sizes to help with quantitative data precision. Surprisingly, it seems that the Blobby shape is the least effective, pleasant, easy-to-use, and suitable shape among all three shape types. This is counter to our original hypothesis that the precise shape would be the least preferred amongst participants. Also of note is that the results we received do tend to be similarly recorded in both the Pattern-Match exercise where we tested how individuals work together to create shapes, and in Swarm-Chase, where individuals react to the crowd shape of a crowd they are not part of. This suggests that crowd shapes are similarly understood from both within and outside the crowd creating it, and that both objectives, working together to form a shape (Pattern-Match) and avoiding a crowd (Swarm-Chase) show no significant difference in participant preference.

This suggests that participants would better enjoy (i.e. find more usable) experiences where they are presented with not merely an abstract shape but also a shape they can recognize as a group of persons. Wanting to see themselves in the shape is not too surprising as Snibbe et al. write of the power of shadows in user experiences [16] and reality-based user interfaces main focusing on “body awareness” [6]. Even in much larger groups our participants seemed much more comfortable with having some sort of visual feedback representing where they are, and where they are relative to others. On the other hand, the limited number of participants and tasks may have affected this finding. We must also note that the groups we tested in the Ball-Catch and Pattern-Match experiments seemed to have known each other. That they are likely more comfortable interacting with each other (i.e. “social embarrassment is less of an issue) could have also skewed these findings. Further studies with strangers may help produce more accurate results though small crowds of strangers could amplify the social embarrassment factor, where larger crowds of strangers may not have as much an effect.

Interestingly, we do see some thoughtful critique on the shape itself, with a large number of comments specifically mentioning how the shape itself could be different. Specifically we see a very large number of comments within the Ball-Catch comment data about the blobby shape that in conjunction with several mentions of frame rate lend some credence to the belief that Ball-Catch did not show any significant results due to technical issues that hampered the participants’ experiences. Because of the many constructive criticisms on the blob shape itself we believe this further strengthens an argument for more study in this area. Even though the blobby shape itself was seen as the

least effective, pleasant, easy-to-use, and suitable, there was definitely a great deal of interest in making the shape better, especially for use within the Combined shape.

We also see a suitable number of comments that found the games themselves interesting, with some interesting comments on how the game mechanics could be “better” – that is, they say, could be changed to provide a more enjoyable and deeper experience. It can be argued that the limited number of experimental tasks and game mechanics developed and researched within this thesis restricts our ability to generalize the findings.

Limited number of participants, the design of test applications, and also the definition of crowd shape (blobby option) can potentially affect the results we have collected. For example, the test applications may be more suitable for one type of crowd visualization, or the design of blobby option may be improved to make it more appealing. Also while the individual feedback seemed important to participants, automated applications may benefit from the blobby option. These lead to the following main research areas that could be further explored to address potential issues with our study:

1. Further studies with more participants and tasks to verify whether the combination shape is a preferred method of interaction over the precise shape.
2. Exploration into what shape properties participants would prefer with focus on variation of blobs and silhouettes.

3. Increasing the group sizes, in real-world contexts, to see if these results would scale up to much larger crowd sizes, at say a concert or outdoor festival; and also allow us larger data sets to explore any gender, educational, or computer expertise differences.
4. Exploration of games and game mechanics, or in general task structures which could be suitable for various visualization methods
5. Investigating automated pattern recognition algorithms that may detect particular crowd features such as energy, mood, etc. from the variations of blobby shape

## 7 Conclusion

This research aimed at answering two main questions regarding the crowd representation in HCI:

- What types of interaction are best suited to crowd visualization?
- How best do we handle visual feedback when dealing with a crowd?

We explored these questions by identifying three visualization methods or crowd shapes (precise, blobby and combined) and we evaluated their performance in two general types of interaction (user as part of the crowd, and user against a crowd). We defined and attempted to verify three hypotheses that would represent our initial assumption about the research questions. They claim that the combined shape is the most successful and the precise shape the least successful in both types of interaction and based on a series of evaluation criteria. Our criteria included effectiveness, pleasantness, ease of use, and general suitability.

While the qualitative results do in fact show a preference for the combined visualization, we noticed that (1) the statistical analysis does not provide strong evidence of its superiority over the precise option, and (2) the blobby visualization is least favoured both quantitatively and qualitatively. The results apply to both types of interactions (Pattern-Match where we observed participants working together and Swarm-Chase where participants individually reacted to a simulated crowd) and suggest that while the overall concept of a single crowd shape is important and helpful, the individual feedback still plays a significant role. This provides a general guideline for

designers creating systems that involve crowds. Such system can benefit from a single crowd shape (variations of the blobby option) but should not remove the individual feedback altogether.

When dealing with automated systems such as crowd control on the other hand, the individual feedback may be less important. This has not been part of our current research and should be considered a direction for further research. Other directions to be considered for further continuations of this project include design and evaluation of other possible crowd shapes and using various test application with more participants, and in more real-world contexts, all of which may affect the conclusions of this research.

## References

1. Barkhuus, L., & Jørgensen, T. (2008, April). Engaging the crowd: studies of audience-performer interaction. In CHI'08 extended abstracts on Human factors in computing systems (pp. 2925-2930). ACM.
2. Brignull, H., & Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. In Proceedings of INTERACT (Vol. 3, pp. 17-24).
3. Brown, B., O'Hara, K., Kindberg, T., & Williams, A. (2009, April). Crowd computer interaction. In CHI'09 Extended Abstracts on Human Factors in Computing Systems (pp. 4755-4758). ACM.
4. Dix, A., & Sas, C. (2008, April). Public displays and private devices: A design space analysis. In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI 2008).
5. Finke, M., Tang, A., Leung, R., & Blackstock, M. (2008, September). Lessons learned: game design for large public displays. In Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts (pp. 26-33). ACM.
6. Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008, April). Reality-based interaction: a framework for post-WIMP interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 201-210). ACM.

7. Kaviani, N., Finke, M., & Lea, R. (2009). Encouraging crowd interaction with large displays using handheld devices. In *Crowd Computer Interaction Workshop at CHI*.
8. Maynes-Aminzade, D., Pausch, R., & Seitz, S. (2002, October). Techniques for interactive audience participation. In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces* (p. 15). IEEE Computer Society.
9. McPhail, C., & Wohlstein, R. T. (1983). Individual and collective behaviors within gatherings, demonstrations, and riots. *Annual Review of Sociology*, 579-600.
10. Murphy, T. (2008, October). Success and failure of audience response systems in the classroom. In *Proceedings of the 36th annual ACM SIGUCCS fall conference: moving mountains, blazing trails* (pp. 33-38). ACM.
11. Nielsen, J. (1994). *Usability engineering*. Elsevier.
12. O'Hara, K., Glancy, M., & Robertshaw, S. (2008, November). Understanding collective play in an urban screen game. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 67-76). ACM.
13. Reeves, S., Benford, S., O'Malley, C., & Fraser, M. (2005, April). Designing the spectator experience. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 741-750). ACM.

14. Saltz, D. Z. (2001). Live media: Interactive technology and theatre. *Theatre Topics*, 11(2), 107-130.
15. Shneiderman, B. (1988). We can design better user interfaces: A review of human-computer interaction styles. *Ergonomics*, 31(5), 699-710.
16. Snibbe, S. S., & Raffle, H. S. (2009, April). Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1447-1456). ACM.
17. Tang, A., Finke, M., Blackstock, M., Leung, R., Deutscher, M., & Lea, R. (2008, April). Designing for bystanders: reflections on building a public digital forum. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 879-882). ACM.
18. Winkler, T. (2000, December). Audience participation and response in movement-sensing installations. In *Proc. of the International Computer Music Conference* (December 2000).
19. Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human-Computer Interaction*, 1(4), 311-338.
20. Vogel, D., & Balakrishnan, R. (2004, October). Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (pp. 137-146). ACM.

21. Hornecker, E., & Buur, J. (2006, April). Getting a grip on tangible interaction: a framework on physical space and social interaction. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 437-446). ACM.
22. Mehran, R., Oyama, A., & Shah, M. (2009, June). Abnormal crowd behavior detection using social force model. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 935-942). IEEE.
23. Rodriguez, M., Laptev, I., Sivic, J., & Audibert, J. Y. (2011, November). Density-aware person detection and tracking in crowds. In Computer Vision (ICCV), 2011 IEEE International Conference on (pp. 2423-2430). IEEE.
24. Davies, A. C., Yin, J. H., & Velastin, S. A. (1995). Crowd monitoring using image processing. *Electronics & Communication Engineering Journal*, 7(1), 37-47.
25. Chan, A. B., Liang, Z. S., & Vasconcelos, N. (2008, June). Privacy preserving crowd monitoring: Counting people without people models or tracking. In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on (pp. 1-7). IEEE.
26. Reynolds, C. W. (1999, March). Steering behaviors for autonomous characters. In Game developers conference (Vol. 1999, pp. 763-782).
27. Gibbons, J. D. & Chakraborti, S. (2011). *Nonparametric statistical inference* (5th ed.). Boca Raton, FL: Chapman & Hall/CRC Press.

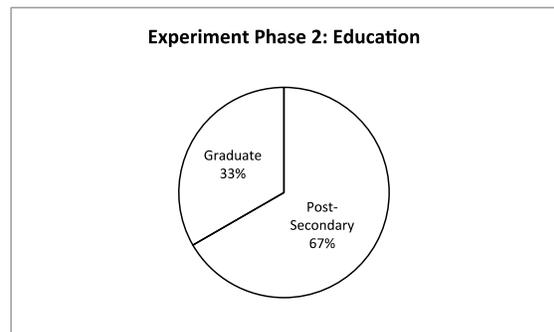
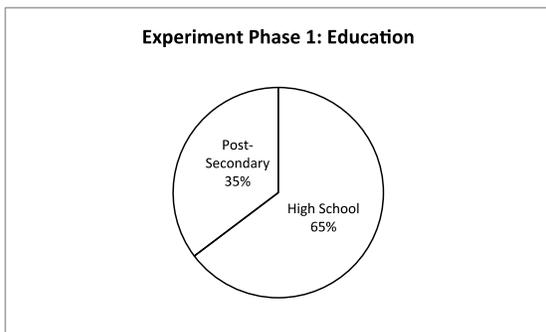
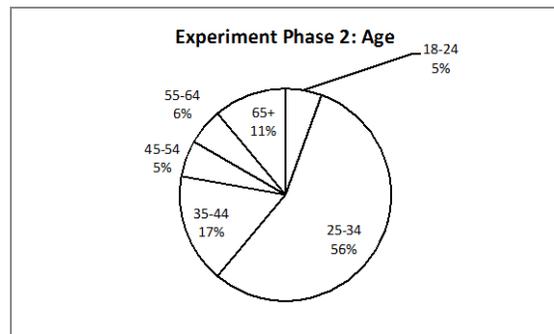
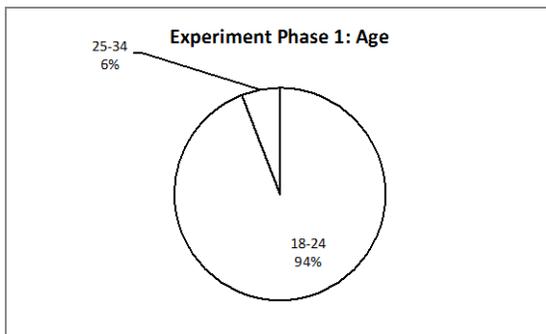
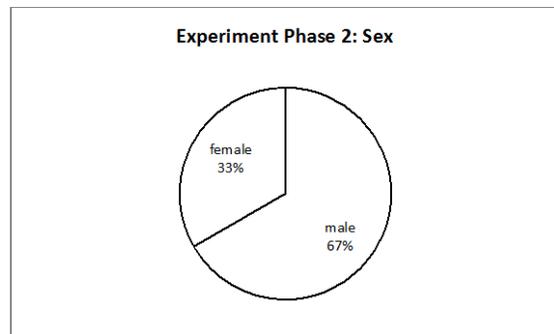
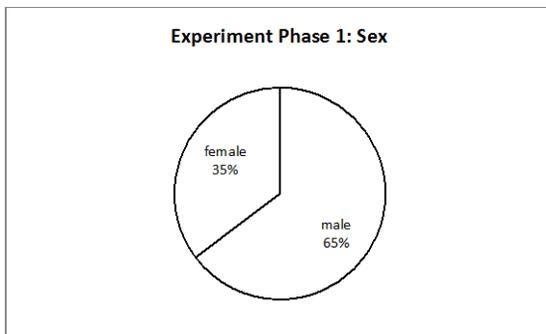
28. Kim, P. (2013, January 7). Awesome Universe of Creative Coding, Explained in Five Minutes. Retrieved January 1, 2015, from <http://createdigitalmotion.com/2013/01/awesome-universe-of-creative-coding-explained-in-five-minutes-video/>
29. Turner, R. H., & Killian, L. M. (1957). Collective behavior.
30. OpenCV. (n.d.). Retrieved January 1, 2015, from <http://opencv.org/>
31. Bourke, P. (1989). Efficient triangulation algorithm suitable for terrain modelling. In Pan Pacific Computer Conference, Beijing, China.
32. Kinect for Windows. (n.d.). Retrieved January 1, 2015, from <http://www.microsoft.com/en-us/kinectforwindows/>
33. McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. *BMC medical research methodology*, 7(1), 30.
34. Albertazzi, L. (Ed.). (1999). *Shapes of Forms: From Gestalt psychology and phenomenology to ontology and mathematics (Vol. 275)*. Springer Science & Business Media.
35. Welch, G., & Bishop, G. (2001). An introduction to the kalman filter. *Proceedings of the Siggraph Course*, Los Angeles.
36. Daniel, W. W. (1990). *Applied nonparametric statistics (2nd ed.)*. Boston, MA: Cengage Learning

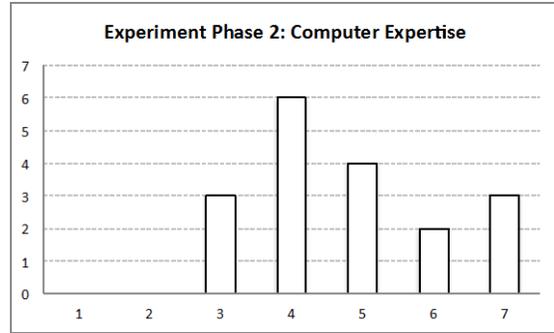
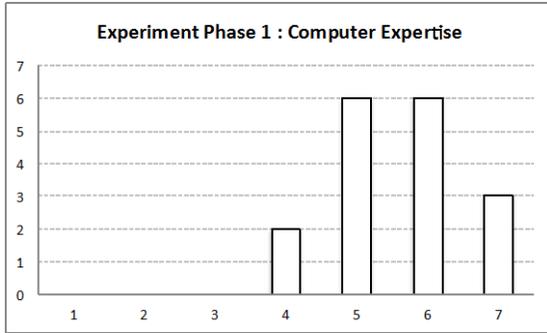
37. Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). Hoboken, NJ: John Wiley & Sons.
38. Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods* (2nd ed.). Hoboken, NJ: Wiley.
39. Sheskin, D. J. (2011). *Handbook of parametric and nonparametric statistical procedures* (5th ed.). Boca Raton, FL: Chapman & Hall/CRC Press.
40. Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2009). Usability: a critical analysis and a taxonomy. *International Journal of Human-Computer Interaction*, 26(1), 53-74.
41. Reeves, S. (2011). *Designing interfaces in public settings: Understanding the role of the spectator in Human-Computer Interaction*. Springer Science & Business Media.

# Appendices

## Appendix I: Demographics

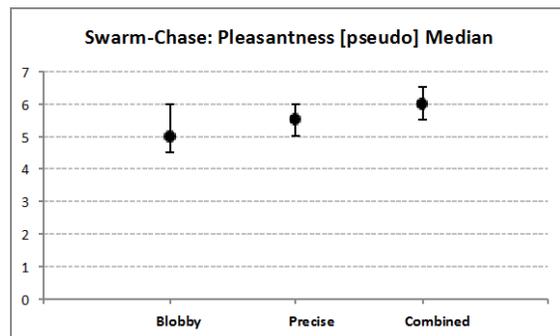
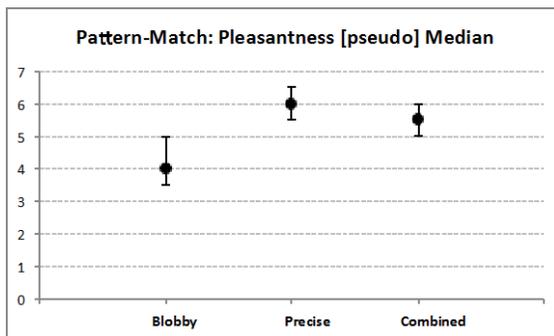
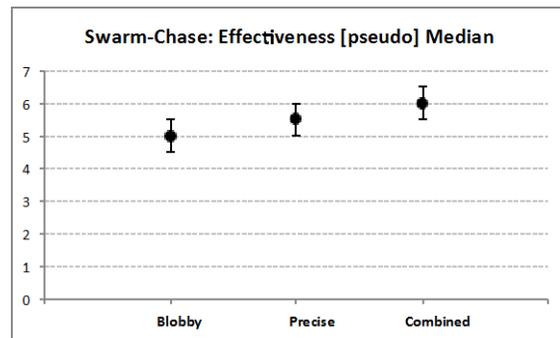
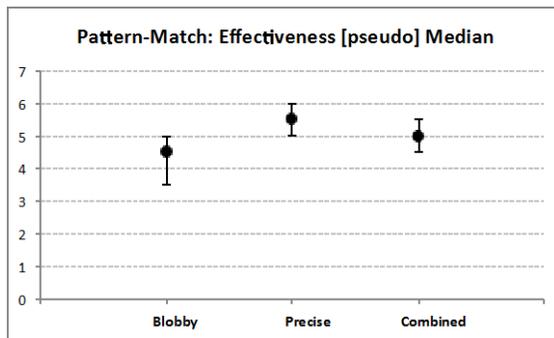
This Appendix contains all demographic information recorded for both experiment phases. Experiment phase 1 (crowd) included participants for the Ball-Catch and Pattern-Match experiments; and phase 2 included participants for the Swarm-Chase experiment.

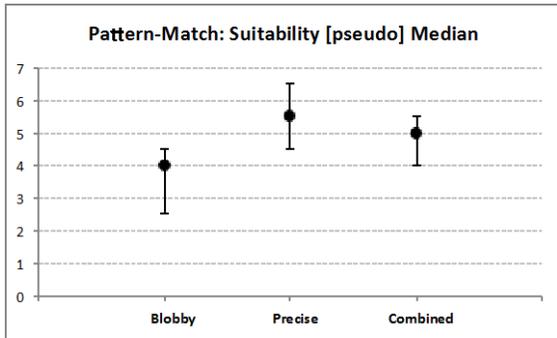
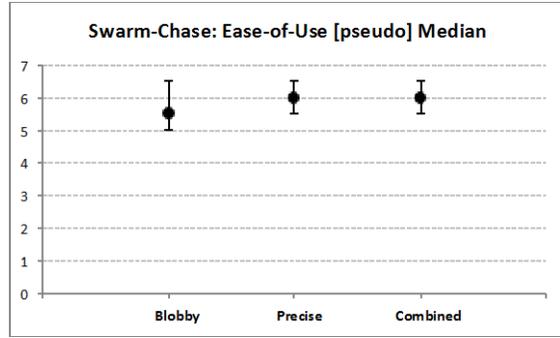
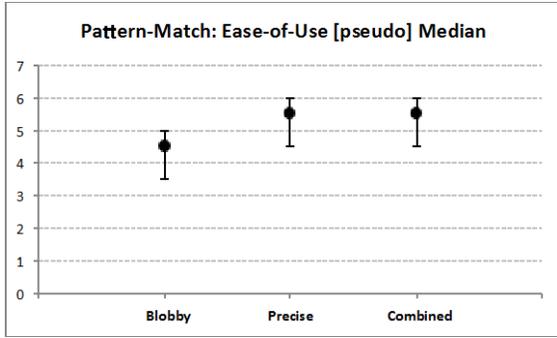




## Appendix 2: Pseudo-Median Point Estimates

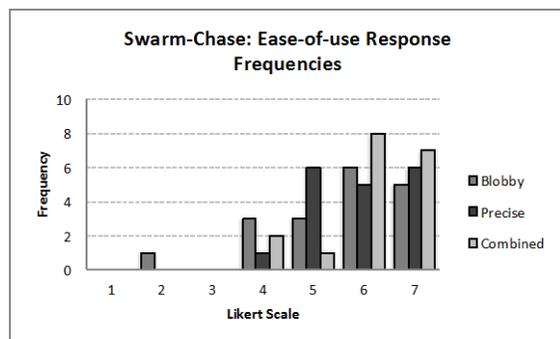
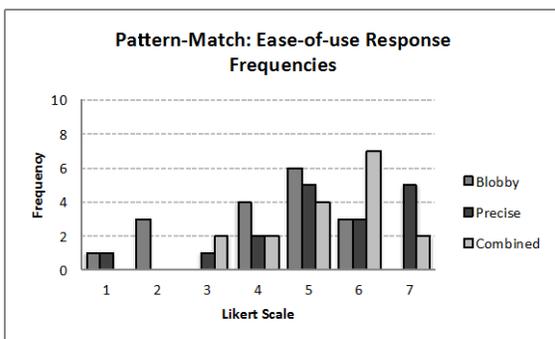
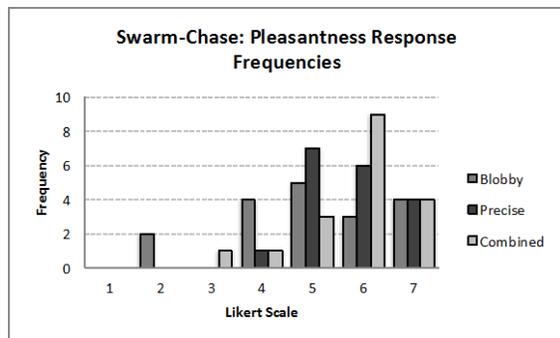
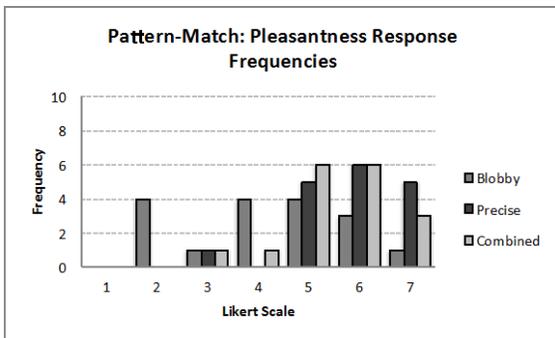
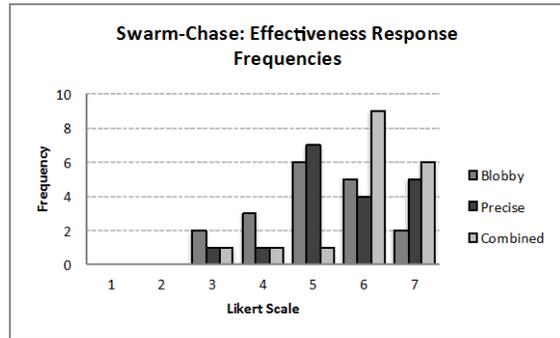
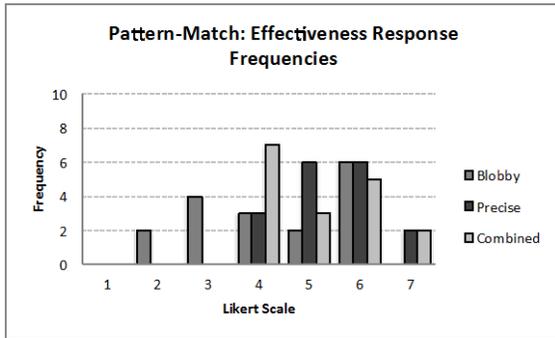
This Appendix contains point estimates (pseudo medians) for all experiments' populations (*Pattern-Match*, and *Swarm-Chase*), across all four dimensions of measure (*Effectiveness*, *Pleasantness*, *Ease-of-Use*, and *Suitability*) with confidence intervals so that we can better visualize the central measure for each. We only graph the point estimates where significance was detected and thus we ignore the *Ball-Catch* experiment, and *Swarm-Chase*'s *Suitability* as there were no significant results detected.

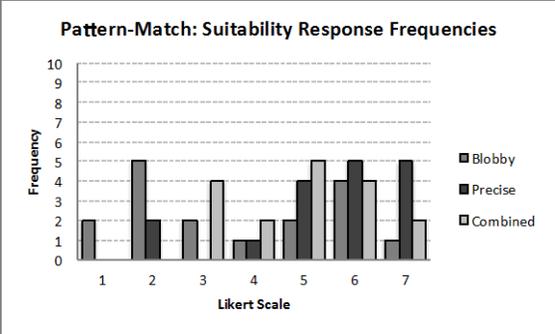




### Appendix 3: Questionnaire Response Data

This appendix displays our questionnaire response frequency data for Pattern-Match and Swarm-Chase. We only graph the dimensions where significance was detected





## Appendix 4: Questionnaire Comment Coding Tables

This appendix displays how we coded and evaluated the open-ended questions of our questionnaires.

**Table 8.** Ball-Catch Comment Data. The four most common themes within the open-ended questions are presented in this table.

	Combined/Both preference	Technical Issues	Blobby shape critique	Game critique
1	"I think this was the most effective method because you get to see how you were controlling the blob"	"Framerate bad for pleasantness. Good minus frame rate"	"Seeing more precision was easier to control the ball"	"Maybe change the colour of the blob so player's can identify with"
2	"This method was the most rewarding, had a lot of control over movement of the shapes and working with others."	"Framerate is bad; but interaction was fun."	"Tighten blob to silhouettes"	"The blob was hard to control."
3	"Guide was helpful, perhaps less blob"	"Low fps eh?"	"make it so the precise shape does not go outside the blobby shape" [Combined shape]	"Maybe read the crowd better it dropped my arms and didn't notice."
4	"Good minus frame rate"	"...but because the program was slow down slightly, it was slightly harder to control"	"Same as blob, not so boxy on the outlines" [Combined Shape]	"Have the balls treat the blob borders as a harder edge"
5	"Liked both better than blobby. I liked seeing the silhouette better makes it easier to understand how to move"		"maybe not as blobby" [Blobby Shape]	"The effect on the ball was not strong enough" [Precise Shape]
6	"... interaction was fun."		"Make shapes more accurate. Hard to distinguish who is where." [Blobby shape]	"The detection should be more accurate. Falling balls should accelerate more."
7			"hard to see yourself and others and yourself to control shapes and see where everyone	"Working together to move the precise shapes were difficult without the bobby

		was" [Blobby Shape]	shapes"
8		"The outline method could be a bit more precise. For example, if I stick out an arm out the protrusion should be narrower"	"It did not seem to react well when we touched the ball." [Precise Shape]
9		"Have the blob a little more constrained to the players so that it is easier to manipulate"	"Have the velocity of the hand motion affect the balls more strongly. It's hard to change ball positions"
10		"Probably allow for more slopes to be formed. It was a of flat areas with curved edges"	"Shapes kind of interfere with each other. Not very clear how to interact the first time." [Combined]
11		"More refinement of the blob shape."	
12		"Make the blob's edges match the silhouette shapes closer"	
13		"Smoother edges .. They looked like rounded boxes which is vague"	

**Table 9.** Pattern-Match Comment Data. The four most common themes within the open-ended questions are presented in this table.

Categories	Combined/Both preference	Technical Issues	Blobby only not viable	Shape to match too complex
1	"Best method for filling in more precise areas ( horns and tail )."	"It was rather strenuous to get into position with the subject so low on the screen"	"maybe make the blobby shape a bit more defined because it was hard to tell who was who when filling the shape" [Blobby Shape]	"Since there is some space between the player and the edge of the blob there was no way to control the precision of the horns or tail"
2	"this method was pretty good but maybe make the blobby shape a little smaller"	"Shape was a bit small for 6 with large blobbiness, slightly closer?"	"Hard to distinguish who is what shape to coordinate"	"Blobby shape is very vague so harder to match more detailed / complicated shapes"
3	"Control the amount of blobbiness. Somehow	"Fill in more space but	"Make blob match detected silhouettes	"Easy to match details but harder to fill a

	to better fit the shape. Lots of fun!"	too little control."	more closely, so shape is easier to manipulate"	volume." [Precise shape]
4	"Was easier to get an approximate shape but the blob outline was too big"	"the precision lead to more artifacts along outline." [Precise Shape]	"We had no idea who was affecting the outline where!"	
5	"Best. Would be nice to have a less wide shape? Most control. Easy to see progress."	"Hard to do small movements"		
6	"Framerate but otherwise excellent!"	"Framerate ..."		

**Table 10.** Swarm-Chase Comment Data. The four most common themes within the open-ended questions are presented in this table.

Categories	Combined/Both preference	Technical issues	Blobby blur critique	"Game" critique
1	"better having the combination of shapes."	"To move up and down was slightly difficult. Hand gestures would have made this easier."	"The blurred edges of the blob make it more pleasant to interact with; but also more vague as to which target to avoid."	"Using simple shapes. For example Circle."
2	"Method was easier"	"Was laggy a bit. Sometimes disappears."	"Gives a good idea of blob shape but little information on what the blob represents or how to interact. Also hard to tell where blob edges are."	"Make the shapes more interesting ( visually stimulating )"
3	"This was my favourite b/c I could see the "people" and the boundaries of the blob + shading"	"feels a bit laggy"	"I think the blur of the shape can be much higher so it feels more organic."	"add sound"
4		"Speed up the tracking"	"Adding visible thresholds - adding stroke"	"have a definite relationship between the crowd and subject. Have subject do more than just avoid crowd."
5		"maybe faster."		"The crowd seems always to move in the same pattern. If there were some "stragglers" that wondered away"

6

“Enhanced facial  
detection or movement  
capture would have  
increased stability.”

from the main group”  
[precise shape]

“They all seemed to  
move at the same pace.  
Maybe if they [people]  
were different speeds”

---

## **Appendix 5: Questionnaire for Ball-Catch and Pattern-Match**

In this appendix we have the questionnaire we used for both the Ball-catch and Pattern-Match exercises, with some slight formatting to save space.

### **Crowd-Computer Interaction Research**

In this survey we will ask some general questions about the participant and provide some background information.

#### **Objective**

In this experiment we will be testing how multiple participants interact with a large screen that displays various visual feedback that displays the crowd in various forms.

#### **Process**

There will be three sections for this experiment, with survey questions for each to fill out. The entire experiment should take approximately 30 minutes to complete.

The basic steps are as follows:

- Researcher gives brief overview of experiment
- Fill out basic personal information on survey
- Complete first phase of experiment “Ball-Catching Game”
- Fill out “Ball-Catching Game” section of survey
- Complete second phase of experiment “Shape-Matching Game”
- Fill out “Shape-Matching Game” section of survey

#### **Personal Information**

1. Age (years)

---

2. Gender

Male Female Other

3. Highest education level attained

Secondary

Post-Secondary

MA

PhD

4. On a scale of 1 to 7, what is your expertise with computers?

1 2 3 4 5 6 7

## 1.0 Ball-Catching Game

---

In this section of the experiment we are gauging how participants played the ball catch game together.

### 1.1 Using Blobby Shape

1. On a scale of 1 to 7, how effective was the interaction method?

1 2 3 4 5 6 7

2. On a scale of 1 to 7, how pleasant was the interaction method?

1 2 3 4 5 6 7

3. On a scale of 1 to 7, how easy to use was the interaction method?

1 2 3 4 5 6 7

4. On a scale of 1 to 7, overall how suitable was

the interaction method?

1      2      3      4      5      6      7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

### 1.2 Using Precise/Coloured Shape

1. On a scale of 1 to 7, how effective was the

interaction method?

1      2      3      4      5      6      7

2. On a scale of 1 to 7, how pleasant was the

interaction method?

1      2      3      4      5      6      7

3. On a scale of 1 to 7, how easy to use was the

interaction method?

1      2      3      4      5      6      7

4. On a scale of 1 to 7, overall how suitable was

the interaction method?

1      2      3      4      5      6      7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

### 1.3 Using Blobby AND Precise/Coloured Shape

1. On a scale of 1 to 7, how effective was the interaction method? 1 2 3 4 5 6 7
2. On a scale of 1 to 7, how pleasant was the interaction method? 1 2 3 4 5 6 7
3. On a scale of 1 to 7, how easy to use was the interaction method? 1 2 3 4 5 6 7
4. On a scale of 1 to 7, overall how suitable was the interaction method? 1 2 3 4 5 6 7
5. Depending on your previous answers, what would you change to make this interaction method better?
- 

## 2.0 Shape-Matching Game

In this section of the experiment we are gauging how participants played the shape match game together.

### 2.1 Using Blobby Shape

1. On a scale of 1 to 7, how effective was the interaction method? 1 2 3 4 5 6 7
2. On a scale of 1 to 7, how pleasant was the interaction method? 1 2 3 4 5 6 7
3. On a scale of 1 to 7, how easy to use was the

interaction method? 1 2 3 4 5 6 7

4. On a scale of 1 to 7, overall how suitable was

the interaction method? 1 2 3 4 5 6 7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

## 2.2 Using Precise/Coloured Shape

1. On a scale of 1 to 7, how effective was the

interaction method? 1 2 3 4 5 6 7

2. On a scale of 1 to 7, how pleasant was the

interaction method? 1 2 3 4 5 6 7

3. On a scale of 1 to 7, how easy to use was the

interaction method? 1 2 3 4 5 6 7

4. On a scale of 1 to 7, overall how suitable was

the interaction method? 1 2 3 4 5 6 7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

## 2.3 Using Blobby AND Precise/Coloured Shape

1. On a scale of 1 to 7, how effective was the interaction method?                    1        2        3        4        5        6        7

2. On a scale of 1 to 7, how pleasant was the interaction method?                    1        2        3        4        5        6        7

3. On a scale of 1 to 7, how easy to use was the interaction method?                    1        2        3        4        5        6        7

4. On a scale of 1 to 7, overall how suitable was the interaction method?                    1        2        3        4        5        6        7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

## Appendix 6: Questionnaire for Swarm-Chase

In this appendix we have the questionnaire we used for the Swarm-Chase exercise, with some slight formatting to save space.

### Crowd-Computer Interaction Research

In this survey we will ask some general questions about the participant and their experience with the prototype interactive installation.

#### Objective

In this experiment we will be testing how a participant interacts with a simulated crowd visualized on a large screen.

#### Process

There will be three sections for this experiment, with survey questions for each to fill out. The entire experiment should take approximately 15 minutes to complete.

The basic steps are as follows:

- Researcher gives brief overview of experiment
- Fill out basic personal information on survey
- Complete the three sections of the survey.

#### Personal Information

1. Age (years)

\_\_\_\_\_

2. Gender

Male    Female    Other

3. Highest education level attained

Secondary

Post-Secondary

MA

PhD

4. On a scale of 1 to 7, what is your  
expertise with computers?

1 2 3 4 5 6 7

## 1.0 Swarm Chase Game

---

In this section of the experiment we are gauging how participants avoid the crowd using three forms of crowd shape display.

### 1.1 Using Blobby Shape



1. On a scale of 1 to 7, how effective was the  
interaction method?

1 2 3 4 5 6 7

2. On a scale of 1 to 7, how pleasant was the interaction method?                    1        2        3        4        5        6        7

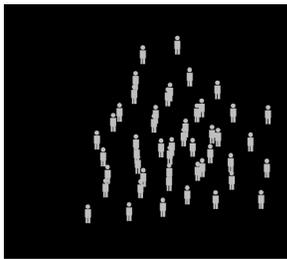
3. On a scale of 1 to 7, how easy to use was the interaction method?                    1        2        3        4        5        6        7

4. On a scale of 1 to 7, overall how suitable was the interaction method?                    1        2        3        4        5        6        7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

### 1.2 Using Precise/Coloured Shape



1. On a scale of 1 to 7, how effective was the interaction method?                    1        2        3        4        5        6        7

2. On a scale of 1 to 7, how pleasant was the interaction method?                    1        2        3        4        5        6        7

3. On a scale of 1 to 7, how easy to use was the interaction method?                    1        2        3        4        5        6        7

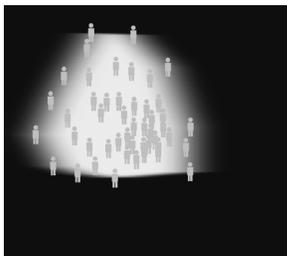
4. On a scale of 1 to 7, overall how suitable was the interaction method?

1      2      3      4      5      6      7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

### 1.3 Using Blobby AND Precise/Coloured Shape



1. On a scale of 1 to 7, how effective was the interaction method?

1      2      3      4      5      6      7

2. On a scale of 1 to 7, how pleasant was the interaction method?

1      2      3      4      5      6      7

3. On a scale of 1 to 7, how easy to use was the interaction method?

1      2      3      4      5      6      7

4. On a scale of 1 to 7, overall how suitable was the interaction method?

1      2      3      4      5      6      7

5. Depending on your previous answers, what would you change to make this interaction method better?

---

## Appendix 7: Ethics Clearance Form



Carleton University Research Office  
Research Ethics Board  
1325 Dunton Tower  
1125 Colonel By Drive  
Ottawa, ON K1S 5B6 Canada  
Tel: 613-520-2517  
ethics@carleton.ca

### Ethics Clearance Form – New Clearance

This is to certify that the Carleton University Research Ethics Board has examined the application for ethical clearance. The REB found the research project to meet appropriate ethical standards as outlined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Human, 2nd edition*, and the *Carleton University Policies and Procedures for the Ethical Conduct of Research*.

**Date of Clearance:** April 10, 2014

**Researcher:** Prof. Ali Arya

**Department:** Faculty of Engineering and Design\Information Technology (School of)

**University:** Carleton University

**Project Number:** 101289

**Alternate File Number (if applicable):**

**Project Title:** Crowd Interaction based on Crowd Shape and Dynamic Content

**Clearance Expires:** August 31, 2015

---

### All researchers are governed by the following conditions:

**Annual Status Report:** You are required to submit an Annual Status Report to either renew clearance or close the file. Failure to submit the Annual Status Report will result in the immediate suspension of the project. Funded projects will have accounts suspended until the report is submitted and approved.

**Changes to the project:** Any changes to the project must be submitted to the Carleton University Research Ethics Board for approval. All changes must be approved prior to the continuance of the research.

**Adverse events:** Should a participant suffer adversely from their participation in the project you are required to report the matter to the Carleton University Research Ethics Board. You must submit a written record of the event and indicate what steps you have taken to resolve the situation.

**Suspension or termination of clearance:** Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.



**Andy Adler**  
Chair, Carleton University Research Ethics Board



**Louise Heslop**  
Vice-Chair, Carleton University Research Ethics Board

## Appendix 8: Email for Call for Participants

Hi everyone,

My name is Anthony Scavarelli and I am a student researcher from Carleton University working under the supervision of Dr. Ali Arya. I am conducting a study looking at different methods to visualize a crowd through a single shape and using that shape as means of interaction with multimedia presentations such as artworks.

Participants will be asked to perform some simple tasks in group and in front of camera and big screen. They will see the shape being generated and use that shape to interact with some audio and visual content. They will then fill out a paper questionnaire relating to these tasks. It should take approximately 30 minutes to complete.

To be eligible, participants should be over 18 years old, fluent in English, with no significant visual or auditory impairment.

Ethical Review: This research has been reviewed and approved by the Carleton University Research Ethics Board

If you are interested in participating in this study, please contact Anthony Scavarelli at [anthony.scavarelli@carleton.ca](mailto:anthony.scavarelli@carleton.ca). Thanks!

Regards,

Anthony Scavarelli

## Appendix 9: Consent Form



Carleton University  
1125 Colonel By Drive  
Tel: 613-520-2600  
[www.carleton.ca](http://www.carleton.ca)

The purpose of this informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. This informed consent must provide sufficient information such that you have the opportunity to determine whether you wish to participate in the study.

The research is being conducted by:

**Dr. Ali Arya**  
School of Information Technology  
Carleton University  
1125 Colonel By Drive  
Ottawa, ON, Canada  
[arya@carleton.ca](mailto:arya@carleton.ca)

**Anthony Scavarelli**  
Human Computer Interaction  
School of Information Technology  
Carleton University  
Ottawa, ON, Canada  
[anthony.scavarelli@carleton.ca](mailto:anthony.scavarelli@carleton.ca)

**Purpose:** This research aims at defining a visualization approach that defines an effective “shape” for the crowd as a whole by identifying regions in the scene where crowd members can be recognized and then use colour mapping to show the crowd density in those regions. We then use this shape as means of understanding crowd features (such as mood or intention) and control a multimedia presentation (such as artworks) using the crowd shape. Various parameters and the effectiveness of this approach are the main subjects of research.

**Task Description:** The participants will be in groups of about 10 people and perform simple movements (changing location) in front of camera to make various simple arrangements (like circle and square shape). They will observe visualizations of these shapes in different methods and also have simple interaction (such as taking a shape to get score and control some multimedia content using those shapes). There will be questions to answer after each task.

**Duration:** The experiment should take approximately 30 minutes.

**Potential Risk/Discomfort:** There will be no psychological or physical risk.

**Dissemination:** The result of this research will be used in Mr. Scavarelli’s thesis and might also be published and/or presented in conferences and/or journals, as well as grant applications.

Minimal risk application form September 2012

8

**Anonymity/Confidentiality:** All collected data will remain anonymous. Any data collected will be encrypted and saved in a secure system with password protection at Carleton University. Only the research team (Ali Arya and Anthony Scavarelli) will have access to this data. It may be used for research projects on a related topic.

**Right to Withdraw:** While the individual answers to survey questions can be removed from experimental results, due to the group nature of all tasks it is not possible to remove a participant's movement data from the group (crowd shapes). Participants can withdraw from the study for any reason and at any time prior to the experiment. We ask the participants to let us know (in person or through email) about their withdrawal 24 hours before the scheduled experiment so we can make sure the experiment can go on with the right number of people. During the experiments, participants can still decide to withdraw but the data related to the already performed tasks will be used by research team. Participants may ask that their individual answers be removed from the survey data up to one week after the experiment.

**Compensation:** You will receive a \$5 Tim Horton's gift card.

This research has been reviewed and cleared by the Carleton University Research Ethics Board (REB) and questions and concerns can be addressed to the REB chair.

**Research Ethics Board:**  
Professor Andy Adler, Chair  
Research Ethics Board  
Carleton University Research Office  
Carleton University  
1125 Colonel By Drive  
Ottawa, Ontario K1S 5B6  
Tel: 613-520-2517 E-mail: ethics@carleton.ca

**Consent:** I have read and understand the above terms of user testing and I understand the conditions of my participation. My signature below indicates that I agree to participate in this experiment.

Participant Name: \_\_\_\_\_  
Participant Signature: \_\_\_\_\_  
  
Researcher Name: \_\_\_\_\_  
Researcher Signature: \_\_\_\_\_  
Date: \_\_\_\_\_