Creating Illusion in Computer Aided Performance

Joe Marshall, BA

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

March 2009

Abstract

This thesis studies the creation of illusion in computer aided performance. Illusion is created here by using deceptions, and a design framework is presented which suggests several different deception strategies which may be useful. The framework has been developed in an iterative process in tandem with the development of 3 real world performances which were used to explore deception strategies.

The first case study presents a system for augmenting juggling performance. The techniques that were developed to control this system demonstrate how deception may become useful even when the core of the performance is not deceptive in any way. This is followed by a magic performance called the Cup Game, which was designed to explicitly test the strategies of deception described in the framework. The final case study is an interactive art installation which presents the illusion of a pet rock that lives in a cage. This demonstrates the usefulness of suspension of disbelief in the creation of illusions. It also demonstrates interesting social effects that are used to strengthen this suspension of disbelief.

The idea of creating the impression of a false situation is inspired particularly by previous HCI work on public interaction. This work demonstrated the usefulness of hiding interface use or computer outputs from some people in a situation. The creation of deliberately ambiguous computer interfaces, which allow for a wider variety of interpretations to be made by the user has also been described. The work here goes beyond these techniques to use technology to actively create false impressions. The techniques used in this process are guided by the work of magic performers, and by psychological studies of how magic performance works.

As well as artistic performance, it is envisaged that this work may prove applicable to more traditional situations. In addition to the framework itself, the development of the case studies has created several useful algorithms which have wider applications. The case studies are also useful guides for those creating performance systems, or other systems where deceptive techniques may be useful.

A web page with videos of the three systems is available at
http://www.mrl.nott.ac.uk/~jqm/thesis

Acknowledgements

I'd like to thank my academic supervisors, Steve Benford and Tony Pridmore (and also Steve Mills who moved to New Zealand), for their support during this PhD. This work wouldn't have been possible without their input, inspiration, opportunities, and encouragement. Being part of the University of Nottingham's Mixed Reality Lab was also a great help and inspiration during this work. Thanks go to everyone in the lab.

Secondly, it wouldn't be possible to do this work without anyone to watch it, so thanks to all the lovely people who have given me opportunities to perform, or have taken part in or watched my performances.

Finally, I owe a lot to my wife Penny Munro, for encouraging me when I decided to give up a well paid job to do a PhD in juggling. Little did she know what she was letting herself in for.

Contents

Chapter 1

Introduction

1.1 Overview

In her survey of recent human computer interaction (HCI) theory, Bødker [15] identifies 3 successive waves of HCI. The first involved the study of individual people in lab settings. The second wave, instead of studying the individual user of an application, studied the entire situation in which the actions took place including multiple groups of people who might be involved in the use of a computer system [6]. In the third wave, study of interaction moved beyond the workplace settings of previous studies into a wide range of new social contexts including live performances in theatres, concert halls, galleries, nightclubs and other artistic settings which are the subject of this thesis.

Interactive performances involve multiple people, some of whom are communicating with a computer system, and others who are observing this interaction. Performance situations are inherently unbalanced, with differing knowledge of the system being used - typically there is a mix of expert performers and inexpert audience members or bystanders. A small but growing area of research is studying the challenges posed by the use of technology in performance. This work attempts to provide designers with tools and frameworks to develop and use systems specifically for performance and performance like situations.

A particular focus of previous research has been on managing awareness, understanding and visibility. This area of work challenges designers to consider what they make visible to the various people involved in a performance, and what they choose to hide from, or not to make available to certain people. For example, previous work has studied the effects of hiding things from the audience to a performance [109], the effect of the differing knowledge of performers, participants, and bystanders on the experience [119], and the use of ambiguity as a design tool [50].

In this thesis, rather than focus on what people are made aware of, the process of making people unaware of things is studied. Making people unaware of things is more complex than a simple removal of visibility; It is in fact a complex interaction between the performance system and those interacting within it which is deserving of further study. A key strategy to achieve this is the creation of illusion using deception.

This thesis presents a framework for analysing and designing deception in interactive performance settings which addresses the questions below:-

- What do we wish people to see and understand about a performance?
- When might we wish to use deception in performance?
- What strategies might be used to achieve this deception, and what social and ethical constraints are there on them?

1.1.1 A Focus on Computer Vision

The primary focus of the practical work in this thesis is on the use of computer vision techniques in performance settings. Computer vision was chosen because it can be used with a very wide range of inputs, so is applicable to many situations. It is also non-contact, which is a useful characteristic when creating interfaces for augmented performance - where the interface is often constrained by movements inherent to the performance itself. Computer vision also has a history as a performance technology, having been used in performances and interactive art work since the 1970s (some of this will be described in Section 1.4).

Vision has a particularly rich scope for the alteration of visibility and hiding; in some situations it lends itself to very visible, expressive body movements, making the link between action and computer response clear; in others hidden cameras may perform tracking unknown to an audience, allowing more intriguing effects to be created.

Focusing on one particular mode of interface during this work has also allowed for a greater level of insight into the technology itself. The practical work has in fact created two innovative new computer vision algorithms, a new way of tracking multiple interacting objects and a method of detecting finger pressure visually without augmenting the surface itself. These support new interaction modes that previously would have been impractical or impossible.

1.1.2 Deception & Illusion

Deception is a term which has some negative connotations. However, it is essentially the correct word for the strategies described in this thesis so is used throughout. The use of a loaded term also highlights the ethically constrained nature of the use of these techniques which it is important to consider.

In most uses of deception in performance we will wish to create illusion. We may do this by deceiving people in various ways. Essentially, a deception is an element of a performance which may be used to help the creation of an illusion. An illusion is the holistic effect on the person being deceived of one or more deceptions. As the subject of an illusion someone has a belief that some particular state of affairs is true, when in fact this is not the case. This terminology is consistent with much of the literature on magic, where one or more deceptions (which may include misdirections, sleights, gimmicks, mentalism etc.) are combined to create a single illusion.

1.1.3 Contribution

This thesis aims to make 4 contributions:

1. It defines a new perspective on interaction for the HCI community; that of considering the use of deception and the creation of illusion in interactive systems. This extends existing notions such as ambiguous design, design for multiple interpretations and design for spectator interfaces and framing.

2. It provides a framework and strategies for designing performances that involve deceptions. This is designed to be practical for use by interaction designers when creating real applications. As performances typically contain more elements than just deception and hiding, it is envisaged the framework may be used in tandem with previous HCI frameworks, design guidelines etc.

3. It provides case studies of real-world designs using these concepts. These are intended to be useful to interaction designers wishing to apply the framework, and to artists wishing to build or use similar systems.

4. The novel computer vision algorithms developed during the process of this work are of general applicability; it is hoped that these will be useful to those developing vision based applications in the future.

1.2 Structure

Chapter 2 of the thesis introduces the surrounding literature relating to awareness in HCI and CSCW, and demonstrates a way of describing performances in terms of the perceptions of the people and systems interacting within the performance. The work discussed in this chapter highlights that in many performance settings, as well as making people aware of things, we may wish to adopt deliberate tactics which make people unaware of some aspects of the situation. In Chapter 3 a framework for the use of deception in performance is presented. This framework is grounded in a combination of HCI & CSCW literature and literature relating to the use of illusion in stage performance. The framework presents a set of strategies for performance deception and examples of ways in which to use them.

The framework development has been carried out in tandem with the development of three performance and interactive art works which each explore different elements of the issues raised by the deception framework. These are described in Chapters 4-6.

Chapter 7 provides a detailed discussion of the deception framework with respect to the systems presented here. This is followed by a discussion of the potential for use of these techniques in public interaction and other ‘performance like’ situations which are becoming increasing common. Finally there is a discussion of the ethics of the use of deception in interactive performance, and in particular how the framing of the deception affects this.

1.3 Development and Methodology

1.3.1 Practical Work and Live Performance

The framework has been developed in parallel with the creation of several practical performance pieces. Each of the the practical pieces has been created in a process involving fast prototyping and early testing of prototypes with non-technical users. This approach is a style of work that is common in performance work, with rehearsals, try-out shows etc. and is also becoming increasingly common in HCI research.

Given the essentially social nature of performance, purely lab based testing would not provide a suitable grounding in the use of performance technology. The Juggling Tracker and the Rock have been used by the author to perform in public settings with external audiences during the formation of the ideas in this thesis. Observations from these performances have been used both to improve future performances, and also to inspire and inform the framework development process. Due to the parallel nature of development, the framework has both been inspired by observation of the performances, and been used as a design tool for some elements of the performances. The balance of framework inspiration and use of the framework was different for each piece; these differences will be described in the next section.

The methodologies used for the devalopment and evaluation of each performance differed to some extent, due to the different nature of each performance and also due to the constraints of the situations in which they were performed. When developing the juggling tracker, input from external audiences was used to aid development of the system, and to design the features available within this general juggling performance system. Observations were made both from experience of performance in the system and observations of others using it. In contrast to this, both Rock and the Cup Game were designed primarily by the author, with alterations made after test performances. The Rock was evaluated in a live performance, with notes taken by the author. Due to constraints of the performance setting for Rock, it was not easy to video record interactions with the system. The Cup Game was designed to be highly recordable and structured in order to allow a large range of factors relating to the success of the performance to be explored.

1.3.2 Relationship between Practical Work and Framework

The practical work chapters each relate in a different way to the theoretical work described in Chapters 2 and 3. These chapters are presented beginning with the most straightforward traditional performance setting, with a stage performer and an audience (Chapter 4), broadening out into a less traditionally structured performance in Chapter 5, and finally in Chapter 6 describing an interactive art piece, without the rigid definition of performer and audience of the other two performances. The following lists the way in which each practical exploration relates to the framework development:

- **Juggling Tracker (Chapter 4)** This chapter describes a system for augmented juggling performance. This was initially an exploration of the ideas of visibility described in chapter 2. However, during the development and deployment of this system deceptive elements of the control of the system became clear. These primarily relate to the way in which a single performer combines juggling, control of system outputs, and alteration of the current system configuration. The juggling tracker was the first piece of work developed, and as such was the original inspiration for the exploration of deception and the creation of this framework.

- **The Cup Game (Chapter 5)** The cup game uses similar tracking technology to the juggling tracker and puts it into an explicitly deceptive format based on a common magical trick known variously as the shell game or the cup game. In this trick, an object is hidden under one of three cups, they are shuffled and the audience has to guess which cup the object is under. The computerised version explores two things. Firstly, the use of computer systems as misdirection, and secondly in a 'beat the computer' version of the game, ways in which people try to fool a computer. The cup game is the most recent development, and has been designed entirely using the deception framework. This performance has been designed as a focused test of into the use of deceptive technologies and the usefulness of the framework. In depth analysis has been made possible by the design of an easily recordable setup. This performance is run with one audience member and one performer, and recorded in various ways for post-analysis.

- **Rock (Chapter 6)** Rock is an interactive art installation based around the illusion of a pet rock which lives in a cage and makes noises when touched. This system is based around one central illusion, that of the rock being alive. Rock creates a situation where the performer and audience are much less well defined, with the person interacting with the rock changing all the time and people transitioning from being bystanders to active participants. This also explores the way in which illusion is created in a social setting where multiple people collaborate to create a particular sense of character in the rock. This was developed in the middle of the framework development process and as such makes use of several of the ideas developed during the creation of the juggling tracker. It is described after the Cup Game due to the more complex social situation in this work.

1.4 Computer Vision Performance

A brief introduction to vision based performance is presented here, with selected performances & artistic installations described which demonstrate differing facets of augmented performance that are supported by computer vision. These are presented in three sections which whilst not exhaustive, together cover many of the uses of computer vision in contemporary art and performance. These are described here to demonstrate the broad artistic genre in which the case studies are situated. Each case study also has a short review of closely related artistic work in the relevant chapter.

Computer vision has also been used in other entertainment genres, most notably gaming and outdoor advertising. These are not covered here as they are not the core focus of the thesis.

Responsive Environments Bruce Nauman created several installations in which technology was used to alter the gallery space containing the installation and to change the way in which the viewer interacted with the space. His corridor installations in particular used both live, pre-recorded and delayed video in order to create interesting effects. For example in Live/Taped Video Corridor [154] a video camera at one end of the corridor is connected to a monitor at the other end. This means that as the viewer walks towards the monitor to see what is on it, they are seeing a view of themself from behind. This link between the viewer and the space blurs the boundary between viewer/performer, in effect making the viewer perform to themself (and others viewing the work). Though in many ways these effects are physically similar to those of the fairground hall of mirrors, these are some of the earliest works to use electronic technology to influence and create interesting interactions without any physical interaction with the viewer.

Computer vision technology uses computers to analyse images captured from cameras; this means that Bruce Nauman's work, whilst creating a responsive effect using video technology, does not employ computer vision technology. Probably the first truly computer vision based work was Myron Kreuger's Videoplace [146]. In this, a person stood in front of a backlit wall and the computer analysed their silhouette. On a projection screen, the silhouette was shown and things happened to the silhouette, for example tiny people were shown climbing up the silhouette.

Many later works have used this 'responsive environment' model, basically creating a space in which movement creates a response typically via either audio, or video projections. Another early work was David Rokeby's Reflexions. This used custom built 8x8 pixel video cameras, which worked as an input to a sound generation system, which responded to the position of people within a room. He further developed this idea of responsive sound environments in his Very Nervous System [157], which was also used for live performance. These were explicitly responsive systems - where the performer is not controlling the system, instead the soundscape is responding to their movements.

A recent example of this kind of work is Degradazione per Sovrapposizione di Corpi, by Salvatore Iaconesi and Oriano Persico. This work projects an image onto the floor of a space. Participants are given a normal household broom and are allowed to sweep the image which changes in places where it is swept over into another image [144]. This provides a surprisingly engaging experience; after a short while the image has been swept different amounts in different places and becomes an intriguing patchwork of multiple images.

Tangibles The term tangible interfaces was popularised by Hiroshi Ishii and Brygg Ullmer [69]. These use physical objects to allow people to interact with a computer systems. For example, the reacTable [72] uses computer vision tracked physical objects on a tabletop to create a real time music synthesiser which is controlled by placing objects on the table, and moving the objects around. The original idea of tangible interfaces was in part inspired by art and design projects including Durrel Bishop's Marble Answering Machine design [140] which used physical marbles to represent telephone messages. Marbles could be put into a particular place to replay the message or put onto a phone to call the person back.

Several artistic projects have used a technology known as augmented reality (AR). AR interfaces use special markers designed to be easily tracked by a computer. The 3d position of markers relative to a camera can be detected, and virtual objects can be overlaid on the camera's view, creating a virtual overlay on top of the real world. For example, the Michelangelo project [108] used a pair of video glasses with a camera on front to make a picture frame seem to display an image which morphed between a painting and a picture of the artist who painted that picture, depending on how close to the picture one looked, giving the experience of looking at one picture, only for it to become another on closer inspection.

Augmenting Existing Performance Activities A very different approach to performance systems is taken in the MIT Dance Space project [121]. This uses a conceptual model of a 'hyper-instrument', to create an installation and performance system. A hyper-instrument uses gestural sensing technology to create interesting interactive instruments, that can be played by users. They are 'hyper' in two aspects, firstly that they allow the user to control expressive elements that would not be under their control with a typical musical instrument, and secondly that they use innovative user interfaces. For example video tracking of dancers is used to create projected visualisations which reflect the dancer's position and movements in 3d space. It was used both for live performances and also as a public installation. These hyper-instruments have much more in common with traditional instruments than the work of Rokeby; The instrument is very much under the control of the user, rather than being designed to create an unpredictable dialogue between the user and the system, as in Rokeby's Very Nervous System. The inter-media dance company Palindrome have also created several dance performances using computer vision. Their Eye-Con [156] system uses infra-red based video tracking in order to track silhouettes of performers. The tracking of performers is used to create visualisations and audio which the performers in turn respond to. They combine this with other technology such as monitoring of heart rate, skin contact, muscle tension etc.

Dance Spaces augment the existing performance activity of dancing; several performers have also created augmented instruments using vision technology. These take existing performance instruments, and augment the activity of using the existing instrument. The Facing the Music system [86] uses a hybrid approach, with a standard electric guitar played through an effects unit which is controlled by a face tracking system. The system detects facial movements such as grimaces and applies different levels of common guitar effects based on the facial expression of the user. Several systems have been created that use visual tracking of conducting batons [21, 98, 70], usually these allow the performer to conduct a virtual orchestra or control the speed and volume of music with the baton.

Another orchestral system was used in the ‘Four Senses’, a set of performances produced by Raewyn Turner, and Tony Brooks [20] in 2002 in Auckland, NZ. These used computer vision in order to augment an orchestra, generating real time visualisations based on the movements of the conductor and orchestra. They also translated audio to smell and touch. The idea of this performance was to create a fully inclusive performance which was available to the deaf, blind, or even those missing both senses.

1.4.1 Thesis Work in Context

The juggling tracker is very much an augmentation of an existing performance activity. It takes juggling, and allows the performer to create a new combined performance of juggling and audio visual accompaniments.

The Cup Game is a tangible interaction device framed within a specially setup performance environment. The activities performed within it include the use of the system as misdirection within a standard trick and the use of the system as a tangible interface, both by audience members who don’t know that they are interacting with a computer, and also by knowing audience members.

The rock fits primarily into the tangible interaction category in that it is using the technology in order to make the rock into a tangible interaction device. However, in the way it is placed in a cage, and set up as an installation for the public to interact with, it has many similarities to the way many responsive environments are designed to create an intriguing and ambiguous interaction.

1.5 Impact

The work described in this thesis has led to several high profile peer-reviewed publications. The technology and implementation of the Rock has been presented at a major international HCI conference [90]. The development of the juggling tracker has been described in both a leading international journal on technology & science [88] and also at a conference in that area [87]. The vision algorithms created for the juggling tracker also led to the publication of a computer vision paper in the British Machine Vision Conference [89].

As well as publications, the work here has led to practical systems which have been presented in several successful performances. The juggling tracker has been performed with and demonstrated in public several times, with an audience of approximately 500 people at the largest performance. The Rock has been installed in various situations, including at the Reactor Digital Arts Festival, where it won an award for best installation. The Cup Game has been performed 27 times, 17 of which were performances to members of the public.

1.6 Conclusion

This thesis explores the use of deceptive techniques such as hiding, misdirection, and other forms of trickery in order to create illusion, a belief that a particular state of affairs is the case when in fact it is not so.

One way of studying this is as an issue of awareness. In particular, these techniques can be seen as methods for removing or altering the awareness of others. The next chapter will explore the way in which awareness has been constructed in CSCW and HCI, with a particular focus on visual awareness due to the computer vision based nature of the practical work described here.

Chapter 2

Awareness, Unawareness and Performance

This chapter describes how awareness is a fundamental issue in human computer interaction (HCI) and computer supported cooperative work (CSCW). Three facets of awareness are presented here, each of which are relevant to a different area of the literature. Firstly, how people are made aware of what a computer is doing. Secondly, how computers are made aware of what a person is doing and what is going on in their surroundings. The third type of awareness is how people are made aware of what other people are doing. Many uses of computers are designed to take a situation and add technology in order to improve someone's awareness in one of these ways. A diagrammatic notation is introduced here which can be used to show the awarenesses in a situation and how technology alters this.

Performance is a particularly interesting interaction situation in which these awarenesses are important. Recent work has studied the distinctive issues that arise in a performance situation. One key idea presented by this work is that awareness is not always desirable. In many performances, some things occurring need to be kept hidden from the audience whilst others will need to be emphasised. In such a situation, as well as creating improved awareness of some things, we may also wish to use technology to actively create unawareness of other things. This process of creating unawareness is core to the rest of the thesis.

2.1 Introduction

In this chapter three key strands of HCI and CSCW work are brought together and framed in terms of the way they address different facets of an awareness problem. Each one is related to the question of what a person (or a computer system) is aware of and how they are made aware of this.

The first type of awareness is the way in which HCI addresses how a computer makes a person aware of what is happening. Secondly, with the use of sensor technology to extend interfaces beyond well understood physical computer interfaces, what the computer is able to sense, and how it is able to sense it has become an interesting focus of research. The final type of awareness discussed is that of computer supported cooperative work, in which computers are used to mediate communication between multiple people, using the computer technology to create awareness of the actions of each of the other people involved.

A notation is presented in this chapter that unifies these 3 types of awareness onto a single diagram, by using a notation of 'perceptions' of the various people and computer systems involved in a situation. This allows for the awarenesses and unawarenesses in a situation to be described succinctly and is used throughout the rest of the thesis.

Much of the prior work focuses on using technology to create new awareness. Recently however, the spread of public and ubiquitous use of computers has led to the realisation that in many cases it is actually desirable to limit the people who have awareness of some interactions, even when they are performed in the same physical space. In co-located situations, where the default is for actions to be performed in view of others, rather than considering the design of an interface by choosing what to make people aware of, we may instead design by choosing what to make people unaware of.

Interactive performance using sensing systems such as computer vision is a particularly interesting area to focus on when exploring these ideas as the combination of the use of sensing interfaces with the inherently multi-person situation of performance means that all three types of awareness are interesting in this situation. Performance is also particularly relevant in that the roles of performers and audience members very often create situations where we wish to perform actions without making others aware of them.

The need to hide things must be balanced against the need to perform things in full view, meaning that using purely physical hiding of things we wish people to not be aware of may not be possible. In such a situation, we require a way to encourage people to be unaware of things which are actually visible to them. The rest of the thesis will discuss methods for creating such illusions.

2.2 Three Types of Awareness

This section describes the three different types of awareness, and how various strands of HCI and CSCW work relate to each of these forms of awareness.

2.2.1 Human Awareness of a Computer

Early HCI primarily involved modelling of a single user interacting with a computer system. Much of this work involved models such as Fitt's law [47] and Keystroke Level Modelling [25] which allowed low level predictions to be made about the relative efficiency of competing interfaces. In terms of modelling user interaction at a higher level, the computer was modelled as a dumb machine which responds to commands that the user gives it. Users performed tasks using these machines. A particularly influential model of how users perform tasks using a machine was presented in Norman's Design of Everyday Things [101]. Of particular relevance to this thesis are his 4 principles of good design, summarised below:

- Visibility - by looking, the user can see the state of the device
- A Good Conceptual Model - The conceptual model of the device is clear
- Good Mappings - The mappings between control and effects are clear
- Feedback - The user gets constant feedback as to the effects of their actions

Looking at these from the point of view of awareness, visibility and feedback are about exactly what a person is aware of about the current state of the computer, and designing the technology in order to make them aware of particular things such as how the interactions have altered the state of the system.

2.2.2 Computer Awareness (of Humans and Context)

With the increasing use of non-GUI forms of interface such as the computer vision based sensing interfaces described in this thesis, the way in which a computer interprets a person's actions and the way a computer interprets sensor data in order to become aware of its surroundings become far more complex. A key concept in this work is that of context awareness.

Early work on context awareness by Buxton [22] suggested that interfaces should integrate 'background' or 'peripheral' knowledge of context into their workings. His example of this is automatic cameras, where rather than needing to have focus and shutter speed set manually, a single 'take a picture' button causes the camera to perform the task. In this way, many things which were previously parts of the human task are guessed by the camera by interpreting the context of the interaction. The background interaction of the camera detecting its context is combined with the explicit foreground interactions to reduce the effort required by the user to take a picture. Several software systems have been developed which take contextual information into account such as the ContextPhone system [104] which runs on a mobile phone, and tries to determine the current context of the person using the phone, for example whether they are at home, or at work.

One key element of many modern context aware systems is the use of sensors in order to do this background interaction, for example the use of tilt sensors in order to automatically detect which way up the screen is on a PDA [62]. Sensors are also a part of many performance systems and this background, responsive behaviour is often used. Rather than be controlled and told explicitly what to do, a system will respond to a performer's movements in some way. For example the Messa di Voce [147] system uses video tracking and audio input in order to track the position of a singer. This allows it to project visualisations of the sound onto a wall so that it seems as if the visualisations are coming out of the singers mouths along with their singing [83]. The computer uses sensing to respond to position and singing directly, rather than at any time being told exactly where to put the visualisations.

The use of sensors and other non-traditional input devices in applications removes many of the conventions inherent in traditional graphical user interfaces. Bellotti et al. [7] describe 5 questions for designers to ask when making interfaces using these new sensors (Table 2.1) Their design process moves away from task

Address: How do I address one (or more) of many possible devices?
Attention: How do I know the system is ready and attending to my actions?
Action: How do I effect a meaningful action, control its extent and possibly specify a target or targets for my action?
Alignment: How do I know the system is doing (has done) the right thing?
Accident: How do I avoid mistakes?

Table 2.1: Five Questions for Designers of Sensing Systems

based approaches of previous work (in particular that of Norman [101]), and suggests that the interaction process be seen as a dialogue between the user and the system, rather than commanding a task.

An extreme version of this dialogue based approach is explicitly taken in the work of early interactive artist David Rokeby. His responsive sound environments create sound based on movement detected with video cameras, which he describes thus:

> *The installation is a complex but quick feedback loop. The feedback is not simply 'negative' or 'positive', inhibitory or reinforcing; the loop is subject to constant transformation as the elements, human and computer, change in response to each other. The two interpenetrate, until the notion of control is lost and the relationship becomes encounter and involvement.*
>
> *...*
>
> *The installation could be described as a sort of instrument that you play with your body but that implies a level of control which I am not particularly interested in. I am interested in creating a complex and resonant relationship between the interactor and the system.*
>
> David Rokeby: Very Nervous System *[157]*

One key factor identified by the 5 questions is that these sensing systems are not able to sense everything and are not completely reliable. The Expected, Sensed, Desired framework [10] describes a way of taking into account the range of actions a person using a system is likely to perform (expected), the range of actions the system can sense (sensed), the range of actions it is desired that they perform (desired), and conversely those which are to be discouraged. A key focus of the ESD framework is in the areas where the expected, sensed or desired actions do not overlap, and how these may be used. One example which is particularly relevant to performance, is the design of 'expressive latitude', where expected but not sensed actions may actually be useful to a user in providing *"opportunities for readjustment, rest, preparation, follow-through, and performance" [10].* The movement of a performer around a musical instrument such as a piano is a good example of this - it does not essentially change the sound yet is an important part of the performance, both for the performer and also the audience [18] (this also relates to the literature on interpretation and ambiguity, which is covered in detail in Chapter 3).

2.2.3 Human Awareness of Humans

As computing became more widespread, the limitations of the low level single user approach of HCI research became clear; theoretical efficiency from systems studied in the lab did not necessarily translate to efficiency in real world contexts, where multiple people would be involved in the use of a system. The birth of computer supported cooperative work in the 1980s aimed to study systems as part of a larger context such as a workplace:-

'*CSCW should be conceived as an endeavor to understand the nature and characteristics of cooperative work with the objective of designing adequate computer-based technologies*' [5].

CSCW studies may involve sociologists, ethnographers, social psychologists, along with designers and technologists in the analysis of the social settings in which technology is used, the ways in which technology aids communication in these settings and the nature of collaboration which occurs both via and around new technologies.

A key theme of CSCW has been the use of technology to enhance awareness between multiple people. The classic example of this is when workers are collaborating in different offices. They may choose to use video-conferencing to have meetings rather than meeting face to face. However, this technology has never become massively popular. According to Egido [42], the reason for this is in part that many of the most important decisions in business are actually taken during informal and incidental conversations in the workplace. 'Always-on' video conferencing systems, such as the VideoWindow [46], and Mixed Reality Architecture [115] systems aim to address this. They allow incidental conversation and interaction, by creating awareness of people in remote sites at all times. Similar awareness mechanisms have been implemented in other applications such as shared text workspaces [40] and instant messaging [57]. This kind of background awareness by humans is in many ways analogous to the background interaction that computers are performing in Buxton's description of a context sensitive system described above [22]. Awareness of self may also be provided by technology. Several systems use bio-sensing for this kind of interaction, for example the pervasive game Heartlands by Active Ingredient [34, 135] creates awareness of the player's heart rate by making it an active part of the game play.

2.2.4 Awarenesses in Performance

Multiple people are involved in an augmented performance situation, as a minimum an audience and some kind of performer. These multiple people are communicating in some way, either via or with the support of a technological system. Thus computer augmented performance systems can be seen as a particular subset of CSCW. This was recognised as early as the first CSCW conference in 1986, where a paper was presented on the use of technology to perform live text and graphics manipulations [81].

As described in the previous chapter, sensing systems such as computer vision lend themselves to performance use. Because of this, the issues of computer awareness described above are also easy to explore through performance related work. The combination of the need to concurrently interact with a computer and other people also adds depth to the issues of single user awareness of a computer. Performance situations allow the exploration of how all three types of awareness interact with each other.

Several digital performance works have used technology in order to enhance awareness of aspects of the performance that are typically unavailable to the audience. Dance company Palindrome's 'e-touching' system detects skin to skin contact between dancers and uses projections in order to make the force and amount of touch visible to the audience. Similarly, technology has been used to aid performer awareness of the audience and even to replace performers entirely.

One example of a project using technology in this way is the HPDJ project [29]. This uses various ways of measuring audience feedback and uses the audience feedback to direct the music being played, to create a satisfying audience experience similar to that provided by a human DJ. In this performance, arguably the computer replaces the human performer in performing the task of measuring audience feedback, taking on what has been described as the primary act of DJing [49].

Whilst some performers and artists have used the internet or phone networks, and created performance works where the performer and audience communicate over delayed communication mediums, the primary focus of this thesis is on live performances carried out in a single venue. Thus, the analysis here is limited to what are known in CSCW as co-located, synchronous systems [43].

As well as the obvious single person controlled systems, where a performer uses a computer to create output which is consumed by an audience, the area of single display groupware [123] is also of interest to artists and performers. This is where a single display may be concurrently interacted with by multiple people, either by using multiple interaction devices, or by using a form of interaction that is able to sense multiple people at once. Many art installations have been created in this way. For example Rafael Lozano Hemmer's Frequency and Volume [150] allowed multiple people to move in front of bright lights with their shadows on a large wall controlling the tuning of a set of radios in front of the wall. Each person could individually control a radio, or they could come together to control the same one.

2.3 A Notation for Describing Awareness

This section takes the above literature and introduces a diagrammatic notation to give an overview of the nature of awareness in a performance situation.

In order to combine these various HCI aspects which relate to performance, the concept of workspace awareness [55] is adapted here. Workspace awareness is a concept commonly used in groupware and CSCW. It describes a way to break down the nature of awareness of actions performed by others into key aspects of awareness that may be provided in a system. These are summarised by Dix et al. [38](p700) as 'what has happened?', and 'how did it happen?'.

In its original use, workspace awareness relates only to the awareness one person has of the actions of another. However here the concept is extended to take account of computer awareness and the awareness of a person of the computer's actions as described above. This treats the computer as being an active participant in the performance; this is because the success of the work in this thesis relies on an understanding not just of how humans in a performance make sense of what is occurring, but also of how computer make sense of a situation. As the case studies shall demonstrate, with sensing systems such as computer vision, the computer's ability to interpret a situation can be complex and nuanced.

Given the thesis is focusing on performance rather than the conventional description of the people communicating in CSCW as 'P' and 'P', or 'A' and 'B', they are explicitly identified as being in the different roles of Audience and Performer. A separate role for the computer system involved in the performance is also described, this is required in order to take account of computer sensing, the third of the awarenesses.

Three concepts are introduced here to describe workspace awareness in a performance:

Visibility This describes whether the person (or computer) is able to sense a particular action. The word visibility is used here to emphasise the focus in this work on vision based interfaces, but this concept does also apply to audio, touch, and other ways of sensing actions. If an action is visible to a person, this means in the language of CSCW awareness that they know 'what has happened'. In the diagrammatic notation, an action is shown as a **bold** word.

Understanding This describes whether the person or computer knows how something happened. For example, they may be able to see both the input to a computer, and the output that it gives, yet not know how the two are linked. Understanding is described in terms of mappings between actions. In the diagram form, these mappings are shown as *italicised* words.

Perceptions The visibility and understanding of a person or computer is called here the perceptions of that person. Visibility of actions, and understanding of mappings between them, may be placed on a diagram. Figure 2.1 shows an example of these diagrams - in this situation, the actions of a performer are only visible to the computer and the performer, this is mapped into a display output, which is visible to the audience as well. The audience cannot see the performer's actions and does not understand the way in which they are mapped to the output. This

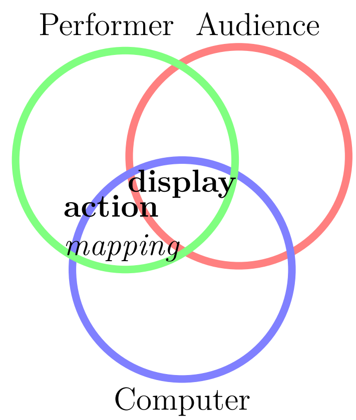


Figure 2.1: An example of perceptions in a performance

notation describes how the perceptions of roles overlap and where they do not. By describing the visibility and understanding in this way several aspects of the performance situation can be described, giving a concise, broad picture description of the aspects of a situation relevant to this thesis. This description brings together the aspects of previous work which are relevant to this thesis - both the awareness work, work on sensing, and the performance and public interaction literature. These diagrams only describe the aspects relevant to this thesis and are not intended to fully describe all of a performance situation. It is worth noting that some existing work describes the interactions occurring in a performance situation in detail (e.g. Dix et al.'s formalisation of performative interaction [39]), but this level of detail is not required here.

Awareness in most CSCW work is about taking things that one person is able to perceivem, such as what they are currently working on, and making them visible to (in the perceptions) of another. There is also the related question, 'how is something happening?', this is basically making underlying mappings between actions visible to the other person. In the remote situations described in most awareness research, actions must be made visible to the computer in order to allow it to make them visible to the second person. Sensing frameworks, which essentially are describing what the computer can see and analyse, can be seen as another facet of awareness, where we are instead looking at what is visible to the computer system, and how it interprets what is occurring.

In terms of performance and public interaction an important aspect of this work is about what the audience / onlookers can or cannot see, how this compares

Question	Example Prior Work
What can the people involved in a performance see and understand?	Workspace awareness (e.g.. [40])
What can the computer systems involved in a performance see and analyse?	5 Questions [7], Expected Sensed & Desired [10], Buxton's Foreground/Background [22] frameworks.
What actions can the performer do which the computer cannot sense?	Expected, Sensed, Desired [10].
What can the computer sense that the audience can or cannot see?	Reeves et al.'s spectator experience framework [109] (see Section 2.4)
What can the computer output that the audience can or cannot see?	Spectator experience framework. [109]
Which links between performers / interfaces in the performance do the audience understand?	Sheridan et al.'s work on performance [119]) (see Section 2.4).

Table 2.2: Aspects of a situation covered by Perception Diagrams

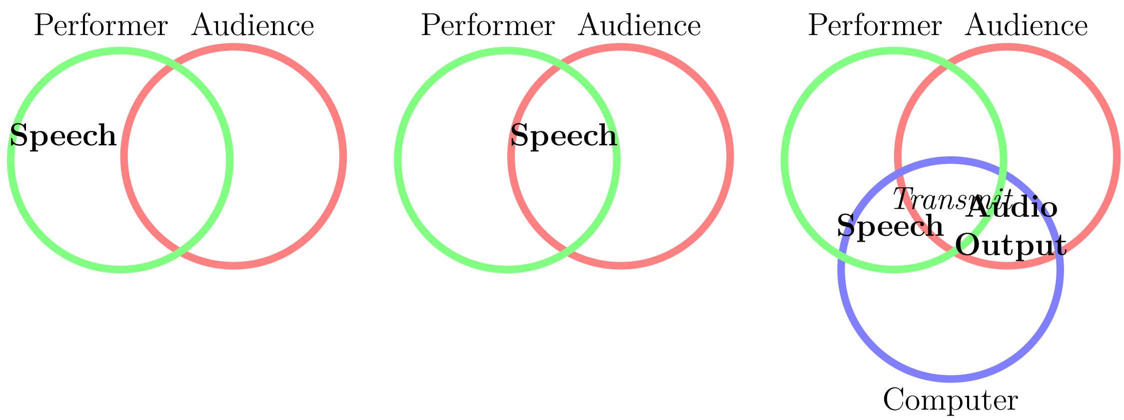
to what is visible to the computer system, and how it is understood by the various people and computer systems involved in a performance situation.

Table 2.2 shows a set of questions describing the aspects of a performance situation covered by perception diagrams, with examples of prior research relating to each question.

2.3.1 Perceptions and Awareness

One way to use perceptions in designing a system is to consider the actual state of awareness and what a desired state might be, and alter the situation to create something analogous to the desired state.

For example in video-conferencing, we have two people, who cannot see each other (Figure 2.2a). The ideal situation would be one in which they can see each other (Figure 2.2b). By using a computer system (actually the entire technological system between both people, in practice a network, plus two computers, cameras + microphones) a situation analogous to the desired situation can be created (Figure



(a)Original (b)Desired (c)Computer Augmented

Figure 2.2: Perceptions in Video Conferencing

2.2c). Note, for simplicity, only one direction of communication is shown in this diagram, clearly in the real situation this is a two way communication.

2.4 Public Interaction and the Value of Unawareness

A distinctive feature of performance as a CSCW situation is that the 'cooperation' aspect is often unbalanced, with performers having a significantly more powerful or active role than audience members or participants in a piece. Reeves et al. [109] identify that the kind of unbalanced situation found in performance is also common to many situations where people interact with computers in public settings. They note how public interaction with computer systems is often performative in nature, where a user of an interface performs their interface use to people looking on. They describe two factors to design interfaces taking this into account - how visible the user's interactions with the system are to onlookers, and how visible the system outputs are to onlookers. They describe the types of interface that may be created by using different combinations of these factors, and reasons why one might wish to use them. Unlike previous work, this does not suggest that awareness is purely positive; several interfaces are described in which hiding the input to, or outputs from the system may be useful, such as 'magical' systems where the outputs are obvious but how they are being produced is not, and 'suspenseful', where the onlooker can see how the person is interacting with a system, yet cannot see how the system is responding until they have a go themselves.

Sheridan et al. [119] describe situations in which audience members become active participants in the performance. Their work relates to 'playful arenas' such as night clubs, where the definition of performer becomes somewhat blurred, as while there are paid performers providing the music and may be other paid entertainers present, much of the entertainment for an individual is created by the other people visiting the club. They describe a model of performances involving three types of people, performers, active participants in a performance, and observers. Key to their work is the awareness that people have of the interactions going on in a performance, for example in the performance described in this paper, the onlooker is only aware of the participant who is carrying a screen attached to their stomach, the performer themselves is using wireless technology to control the screen, meaning that they remain unknown to the onlooker. Again, parts of the performance are deliberately hidden from onlookers to create particular effects.

2.4.1 Perceptions and Unawareness

This section shows in terms of perceptions two examples of how situations with reduced awareness may actually be desirable.

Hyper-Instruments vs. Responsive Spaces

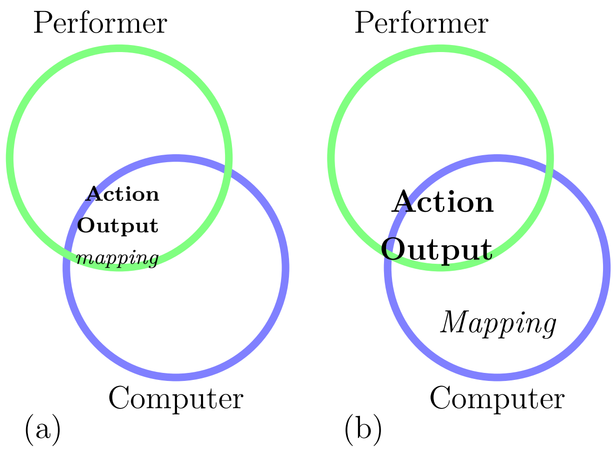


Figure 2.3: (a)Hyper Instruments vs (b)Responsive Spaces

Hyper-instruments, a term originally coined by Tod Machover at MIT, refers to new computer augmented interfaces for musical performance. These take their model from traditional instruments and are designed to be controlled by performers to make particular sounds (or in some cases to produce particular visual outputs). This means that they produce a well understood output when given a particular input and the mapping between the input and output is well defined and can be described to a performer so they can understand the use of the instrument. In this situation, we describe both the performer's actions, the computers output, and the mapping between the two as being perceived by both the computer and the performer (Figure 2.3(a)).

In contrast, responsive spaces may be more exploratory environments, which create a response to the movements of performers or users within them. Whilst they may use similar interfaces such as computer vision based full body tracking, the way in which they respond to these inputs is very different. Rather than the mapping between input and output being clear to the user, the mapping between them may be far less clear, so users are not aware of exactly how it works. This means that they are not able to play the space like an instrument, instead the space responds to them and they must in turn respond to the space, creating a feedback loop. The performer no longer understands the mapping between their input and output, giving the configuration shown in Figure 2.3(b).

Magical vs Suspenseful Interaction

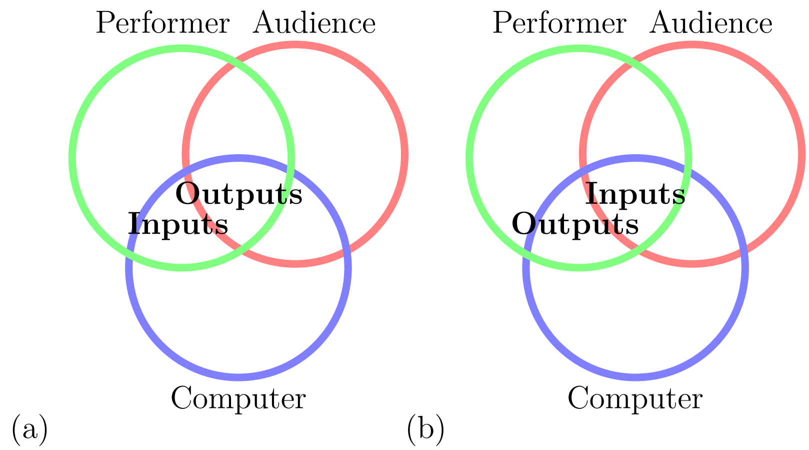


Figure 2.4: (a)Magical vs (b)Suspenseful Interaction

Reeves et al. in their work designing spectator interfaces [109] describe two interesting types of design which have reduced awareness. In their magical interface, the system outputs (the 'effects' of interface manipulation) are revealed, and the inputs (the manipulations of the interface) are hidden from the audience member. This is shown in terms of perceptions in Figure 2.4(a). Their suspenseful interface reverses this, inputs are visible, but outputs are hidden (Figure 2.4(b)).

2.4.2 Actively Creating Unawareness

Most CSCW work has concentrated primarily in situations where lack of awareness is identified as a limiting factor. These solutions create awareness of things which people are not normally aware of. Awareness in this case is about taking a situation where some action *is not visible* to one or more people, and altering the situation into one where it *is visible* to them.

However, in many co-located performance situations the default is actually for things such as the performer's actions to be visible and for audience members to be aware of them. From systems such as Rokeby's responsive artworks and the work on designing for public interaction it is clear that as well as creating new awareness, removing awareness or understanding is also a useful strategy in performance situations.

An alternative way to look at design of perceptions is that rather than being a task of making audience members aware of certain things, it is instead about removing the attention and awareness of audience members from extraneous things. In effect, this means we are taking a situation where an action by default *is visible*, and altering it into a situation where in effect the action *is not visible* to them. Similarly, in terms of understanding, due to the deterministic nature of computers, the default position is that of understandable control, leading artists to deliberately obfuscate systems to make them behave in a less controlled fashion.

Designing in this way can be done by using an approach analagous to the approach shown in Section 2.3.1. First the situation is analysed in terms of real and desired perceptions, then a system is designed which removes awareness.

The key question raised here is how exactly to remove awareness. When considering what to make visible, we have so far considered simply whether we have made it physically available to someone or not. When we take the opposite approach, of trying to hide things that it is inessential or undesirable for someone to see, there are more subtle ways than purely removing it from view. Reeves et al. [109] mention the possibility of partially hiding outputs which is clearly relevant in performance situations, where things as simple as distance from a stage may mean that things are physically visible to someone but are partially hidden by the fact they are too far away to see in detail. They also mention hiding interface manipulations and outputs by performing movements which deceive the onlooker into thinking they are part of other movements but do not discuss in depth how this might be achieved. In this situation, when we are considering a subtractive approach to awareness, visibility and awareness may diverge, creating a situation where despite things being *visible* to a person, they are not *aware* of them.

2.5 Conclusions

This chapter has described how the study of awareness in various forms is a way to understand several key areas of HCI and CSCW. More recently, rather than using technology to create awareness, technology has been employed to keep parts of a multi-person interaction situation hidden from some of the people involved. This approach is particularly suited to performance, where there are often parts of the performer's actions which we may wish to make people unaware of, and other actions which we may wish to emphasise.

This thesis aims to go beyond the previous work by exploring rather than purely physical methods of hiding things from people (such as having displays hidden within booths), methods of hiding things or tricking people which do not require physical hiding. This is useful as in many situations there is no easy way to use physical hiding, particularly if it is desired that other, visible actions are performed concurrently or close in time to the hidden actions.

Prior work from several areas is important in order to achieve this active creation of unawareness. There is some HCI work on ambiguity, interpretation and orchestration of performance which is relevant to this idea. Also, there is a profession that uses strategies of hiding and creating unawareness, that of performance magic. The techniques of magicians have been studied by various authors and a rich literature exists relating to these techniques. A particular magical strategy of interest to this PhD is the use of deception and deceptive framing in order to create illusions and reduce awareness.

The next chapter frames augmented performance with respect to this concept of creating unawareness, and discusses ways in which the techniques of stage magic may be used in order to remove, partially remove or alter people's awareness of part of a situation, to create unawareness of things that are happening, or in order to support an illusion of something that is not actually occurring.

Chapter 3

Altering Awareness Using Illusion

Performers often seem to perform everything in full view of the audience. However, in many cases it is actually desirable to hide certain aspects of the situation from the audience. Because the performer typically needs to be visible to the audience it may not be possible to physically hide performance elements from view, and more subtle deception may be required in order to make people unaware of things.

The HCI literature relating to ambiguity and multiple interpretations is clearly relevant to the task of removing or altering people's awareness of a situation. Similar effects are also used in conventional performance, so it is important to study how it is done there. Stage illusionists, with their use of misdirection and similar techniques provide a particularly explicit example of deceptive techniques. Several studies have been made of the methods of magicians, both by magicians themselves, and also by psychologists interested in the ways magicians exploit the limitations of the human mind.

This chapter draws on a combination of HCI work and studies of magic and illusionism, in order to develop a framework for the use of deception. This framework is based on the cooperative work diagrams introduced by Dix [37] and describes a set of potential strategies for deception, each of which are used for a different type of desired illusion.

3.1 Illusion in HCI

In this section the HCI work relating to illusion is described. This is split into three parts. Firstly, the design for ambiguity and multiple interpretations in HCI is described, this is followed by work relating to human orchestration, where humans are (often hidden) helpers in a computer based performance. Finally, situations where theatre and performance are used as analogies for computer interaction are described.

Ambiguity and Multiple Interpretations Ambiguity can be seen as a way of creating partial awareness; if something is left ambiguous it may be visible, yet not obvious to the person seeing it.

The discussion of ambiguity in design is a relatively recent HCI trend. Gaver, Beaver & Benford [50] suggested that ambiguity, whilst typically seen as a bad thing in interface design, could actually create richer and more interesting interfaces. They reference several artistic systems and interventions and describe the positive nature of ambiguity in terms of encouraging people to engage with interfaces in a particularly personal way, beyond purely that which is envisaged by the original designer.

> *The most important benefit of ambiguity, however, is the ability it gives designers to suggest issues and perspectives for consideration without imposing solutions [50].*

Aoki and Woodruff [4] describe ways in which this analysis extends to communication systems such as email, instant messaging and telephoning. They describe how non face-to-face communication methods add a level of ambiguity, and how this may be socially useful; for example a long pause in an instant messenger conversation may be socially acceptable in a way that a pause in face to face conversation would not be, as it is easy to explain that the person was interrupted or away from their desk. They suggest that this kind of ambiguity should actively be designed into communication systems, as it allows for the stories which we construct in order to maintain social relationships.

Both of these works dealt explicitly with the creation of systems which encourage multiple and flexible interpretations. Boehner & Hancock [16] extend this by looking at the issue of interpretation in a social and temporal context. They describe the way in which we assess the intentions of others by analysing their actions with respect to our interpretation and updating our idea of their intentions accordingly. They describe how over time the multiple users engaged in a system develop conventions and narrow the space of stories around the system as they fit it into their particular social context. Sengers and Gaver [116] also address interpretation and ways in which to explicitly encourage multiple interpretations in design such as forcing users to create their own interpretations by deliberately blocking the most obvious interpretations of the purpose or use of an interface. They argue that multiple interpretations should be considered both in the design and also in the evaluation stage of projects:

> *Design shifts from deciding on and communicating an interpretation to supporting and intervening in the processes of designer, system, user, and community meaning-making. Evaluation shifts from determining whether an authoritative interpretation was successfully communicated to identifying, coordinating, stimulating, and analysing processes of interpretation in practice. [116]*

Several other recent papers have discussed artifacts built using these principles [44, 53, 51, 92] which are designed to support play, interpretation, reflection, and various similar modes of use. One key mode in the design of these artifacts is the process of conversation between designers and users. Ambiguity is argued to be a key part of this as it reduces the privilege of the designer to control the interpretation of the artifact, and forces them to better empathise with the users in the design process [134].

One design that has taken a very concrete approach to multiple interpretations is The Swarm [114], a social networking and messaging tool which explicitly supports the creation of multiple concurrent personas, for example a work and a leisure persona, which are given different information. ‘Dystopian and Utopian’ examples are given on its use, for example allowing a user to share more personal information only with close friends thus maintaining an appropriate divide between their work and personal lives, or (in the dystopian analysis) allowing users to fool their bosses into thinking that they are still at work when they have actually gone out shopping.

These approaches differ from the earlier HCI works engaging with illusion in that they move away from the designer creating a single illusion of what is occurring in the computer, into a situation where the designer is less prescriptive about the illusion they are trying to create and is allowing people to create illusions of their own.

Wizard of Oz & Human Orchestration As well as allowing users the creation of their own illusions, HCI has also explored situations where the illusions are being created by one set of users for others.

The "Wizard of Oz" or WOz method involves a human performing part of a process which is seemingly being performed by a computer. It was first introduced by Kelley [73] as a way of performing iterative design of natural language systems, with earlier versions of the system relying on humans to interpret the user input and mediate the computer's output and increased automation throughout the design process. In its original form, WOz was primarily a prototyping tool. However, hidden human orchestration has been part of several live systems, for example AR Facade [41] uses a hidden person in order to perform gesture and speech recognition, with the human providing a level of interpretation beyond that possible with current computer technology. In the multiplayer mixed reality game Desert Rain [78], the production team communicated to help players using special 'dramatic voices', in order to keep their help embedded within the game. They were also able to secretly nudge participants in the virtual world, for example to get them into the right place if they got stuck in a virtual world. This was done carefully and in very small amounts, in order to be unnoticeable to the players, and to make it seem as if the game was always performing their intentions. As well as simply fighting technical glitches however, this orchestration is described as being used for game related effects, such as to slow down one of the players if they are too far ahead of the other players.

Computers as Theatre Brenda Laurel's 1991 book 'Computers as Theatre' [82] argued for a new approach to human computer interaction, treating computer use as a performance. She relates theatrical theory to the design of computer interfaces, arguing that computer use is less like 'narrative', and more like 'drama', where drama is something which is enacted rather than read. This in in contrast to many design methods involving 'use cases' or other purely narrative descriptions of computer use. Tognazzini [128] is more specific, and describes a set of principles for user interface design based on stage magic, arguing that interfaces are an illusion very similar to stage magic performances that hide the internal workings of the computer and instead present something to the user that is consistent, yet basically untrue. His principles suggest ways to make these illusions work which are directly translated from principles of stage magic as described by Fitzkee [48].

For example, he argues that the stage magician's tendency to manipulate time, so that the actual trick is occurring at a point when the audience thinks something else is going on is similar to a computer program pre-fetching some data when it is anticipated it will be needed soon, so that it can instantly be produced at the desired time.

3.2 The Psychology of Magic

As well as Tognazzini's [128] HCI work based on magic, there is a rich literature directly relating to the techniques of magicians. This is written by psychologists and magicians wishing to understand how illusion works in the particular context of stage magic. Magicians are the most advanced users of deception for artistic purposes, and as such, it is useful to understand their methods.

This section summarises important works and key ideas in this literature that inform this thesis.

Theory of Magic Several early works on magic & deception were written from 1880-1905, at a time when stage magic as performed today was a new and highly popular art form. Many developments were introduced by the famous French magician Jean Eugène Robert-Houdin, who is widely credited as the creator of the modern style of stage magic, with a performer dressed in smart clothes, working in theatres, rather than the street and fairs where magic was previously performed.

Dessoir [36] and Triplett [129] write about the history of 'legerdemain' in society. They describe a rough transition from magic as being performed by priests and shamans and perceived as real sorcery, towards the more modern, professionalised form of staged magic performance shown to an audience who know that they are being tricked.

Several psychological devices magicians use are described in these articles which are fundamental to magic performance in a much more significant way than the raw physical description of how the trick is performed [36, 129]. Jastrow [71] builds upon this work to describe possible psychological processes involved in this deception. Magic is described by these authors as primarily being about leading the audience to see a set of events as a natural progression, only for the events to lead to an impossible conclusion:

> *from this illogical concurrence of two self-contradictory ideas rises the agreeable consciousness of illusion [36](p3609)*

Several methods to achieve this effect are described such as building up a sense of mystery through small simple tricks in order to give the appearance of having higher powers [36], hiding the fact that they are performing dexterous movements [129], the use of gaze and gesture to make unimportant gestures seem important and to draw the attention away from gestures that they are not wishing the audience to see [129]. They place particular emphasis on the importance of the spoken element of the performance or patter in making clear the external story of the trick, and hiding the secret:

> *he says what he does not do, he does not do what he says, and what he actually does he takes particular care not to say anything about [129](p482).*

As one example of this, Jastrow cites Robert-Houdin's advice that if the performer counts "one", "two", "three" in the disappearance of an object, the actual disappearance should have occurred before the performer says "three" as the audience will pay little attention to what happens on "one" and "two" [71](p121).

Repetition is also an important focus in terms of first actually doing what you are pretending to do before performing what is purported to be the same thing as part of the trick [36, 71]. For similar reasons, a rule forbidding repetition of an actual trick is also described as being fundamental to most magicians' practice as much of the power of a trick is formed by the audience not knowing at what point the trick occurs.

Two more recent works apply modern psychological theories to similar issues. Hyman [66] analyses the psychology of deception as a whole, with an emphasis on magic and psychic deception. He suggests that much of the early work described above is surprisingly consistent with today's cognitive psychology. He describes 3 other categories of work relating to deception: 'Non Human Deception' studying how animals deceive and whether animals may intentionally deceive, 'Children and Deception', ways in which developmental psychologists analysed children's development of an ability to deceive, and finally 'Psychic Fraud', where fake psychics have fooled researchers, and psychological mechanisms that allow this to occur.

Hyman also discusses types of deception which occur in some cases as purposeful deceptions and also as self-deceptions, where the person performing the deception truly believes they have some special power. Examples of these include 'cold reading' and spoon bending, both are activities where some 'psychics' delude themselves that they are achieving these effects through mystical powers, whereas others deliberately use deception to convince others of their powers.

Nardi [99] applies a social psychological approach and analyses the situation of a conjuring performance in terms of social interactions, with reference to Goffman's theories of social organisation [54]. He suggests multiple 'channels' of activity occurring during a magic performance. The 'main track' is the effect that the spectator perceives as occurring, for example a card disappearing and reappearing somewhere else. The 'concealed track' is the real sequence of events that makes the trick work. Three other channels are combined in order to achieve the desired effect; 'concealment tracks' physically hide actions from the audience, for example by having objects behind a curtain; 'directional tracks' are verbal and physical cues which direct attention towards events that reinforce the main track, whilst 'disattend tracks' are used to direct audience attention away from parts of the concealed track. He discusses the way in which magicians 'bracket' tricks in time in the different channels, for example with elements of the concealed track for a particular trick occurring at a point where the performer is between tricks or performing a previous trick, whilst the main trick is begun by his actions somewhat later. Nardi also discusses what happens when a trick goes wrong, for example a spectator does not choose a forced card – in this case, Nardi describes the concealed track as being broken, and the performer needing to move onto a different track. However, as the spectator does not know how the trick is meant to happen in many cases the performer can simply change trick to one which does not require the particular forced card. This analysis gives an explanation of the workings behind a method described by Triplett in 1900 [129].

Kuhn, Amlani and Rensink [79] also describe various mechanisms used by magicians - broken down by them into: misdirecting someone, forcing someone to do something, and use of illusions. They call for the creation of a 'science of magic', using knowledge of magical techniques and psychological models in order to explain and categorise the mechanisms involved in magic. They suggest that the exploration of these techniques may create results useful in other fields, including human computer interaction (as suggested by Tognazzini [128]).

Experimental Work on Misdirection Experimental studies of misdirection allow in depth exploration of mechanisms used by illusionists (described in the previous sections) and give insight into the way these mechanisms work. Several of the earlier works described experiments. Recently some experimental psychology research has been done into the mechanisms behind sleight of hand and misdirection.

Triplett describes experimental work where hallucinations are induced into subjects by what is seemingly repetition but in fact is not. One experiment used a wire, heated by passing an electrical current through it. After several trials where the subject held the wire and then it was heated, on the final trial, with no electricity applied to the wire the subjects were in almost all cases still able to perceive a heating up. The application of this idea to conjuring is described in a trick where a ball is thrown into the air three times and is seen to disappear in mid air. In practice, by the third throw the ball has already been dropped, yet viewers present varying suggestions as to at what point the ball disappeared, with some even believing it got as far as the ceiling before vanishing. Binet [13] also describes these illusions. He points out that a key part of the success of these illusions in the hands of magicians is that the audience '*give[s] themselves up to the illusion, without troubling themselves to investigate it*' [13](p557). He argues that the cooperation of the public and the inherent knowledge that they are being tricked is key to many illusions succeeding. This need for suspension of disbelief is clear in other forms of performance such as theatre but less obvious in magic.

More recently, Tatler and Kuhn used eye movement tracking in order to show how magicians use their direction of gaze to stop the audience perceiving their actions, even when the actions are in full view of the audience [127]. This exploits the social tendency to look where another person is looking. They describe this misdirection in terms of a deliberate creation of 'attentional blindness' [120] – a phenomenon where people cannot see an occurrence despite it being fully in their view. They also demonstrate how when the same trick is performed twice, the misdirection fails, as the audience are expecting the trick, and know when to look to catch the magician out, providing an experimental basis for the magical rule of never performing a trick twice (the rule itself is described in most of this literature, eg. [23, 36, 71, 99, 129]). In a related experiment, Wiseman and Greening demonstrate a positive illusion, where adding verbal suggestion to a situation made participants see events occurring that did not happen at all – in this case a spoon bending while no-one is actually touching it [133].

Magicians versus Spiritualists The transition from hidden, mystical trickery to open trickery for entertainment described by Dessoir and Triplett only explains some uses of magic. In the 1890s spiritualism was still popular, and all the early works make particular reference to ways of achieving the magical effects claimed by spiritualist mediums. According to Dessoir, these differ from the effects of magicians primarily in the framing of the experience [36]. A particular emphasis is on the role ambiguity has in making the experience seem paranormal and personalised in some way, for example the same nondescript puppet is reported as being described in séances by different people as being 'grandmother', 'my sweet Betty', 'Papa', 'little Rob' [36](p3664).

Since the emergence of magic as an entertainment, magicians have aimed to expose those who used tricks in order to present themselves as having paranormal abilities, and there has been significant tension between entertainment magicians and those who believe in mysticism. One of the earliest examples is S.J. Davey's séances in 1884 [33]; he used conjuring techniques to recreate séances to demonstrate how mediums could create such effects. Davey's exposure of the tricks of false mediums was not believed and led to him being accused of being a true medium himself [131] even after publishing full details of how he had tricked his audiences. Similarly, the early 20th Century escapologist and magician Harry Houdini dedicated significant parts of his life to recreating and exposing the tricks of false mediums [64], only to be accused by Arthur Conan Doyle of using teleportation to perform his escape from locked containers [30] (the actual methods used by Houdini are described by Cannell [23] in a book released shortly after Conan Doyle's accusation). This tension between stage magicians and spiritualists has continued into the present, most notably in the very public feud between magician James Randi and the mystic Uri Geller [107].

The continuing power of belief in the paranormal was demonstrated recently by Hergovich [59]. He explored experimentally the effect of belief in the paranormal on audience perception of mentalism or mind reading magic tricks. He performed a mind reading magic trick on two audiences, one audience being told that it was an example of psychic powers, the other that it was a magic trick. The resulting study demonstrated that believers in the paranormal would think that paranormal powers were being used even when they were explicitly told that it was a magic trick.

3.3 Key Points from the Psychology of Magic

This section describes several key points from the literature relating to the psychology of magic.

Misdirection is not purely hiding actions The early works on magic all stress that misdirection does not purely mean hiding ones actions visually. In fact, in many tricks the underlying trick is actually occurring in full view of the audience. Tatler & Kuhn verified this particular effect experimentally [127]. As well as *not seeing* things that *are* happening, Triplett [129] and Jastrow [71] both describe the trick where a ball is thrown in the air and seen by the audience to disappear in mid air when in reality the ball has not been thrown up at all, creating a positive illusion of an action occurring. Wiseman and Greening's spoon bending trick [133] provides a further experimental demonstration of this ability to convince the audience that they *are* seeing things which *are not* happening.

Multiple Narratives All the authors above are agreed that the magician is keeping track of multiple narratives, both the external narrative as perceived by the audience, and the internal workings of the trick. The importance of the visual presentation of the external narrative is shown by the recommendation of the practicing of tricks in a mirror prior to showing them to any real audience.

Time Framing The internal and external narratives may not be synchronised in time – for example setup for a trick may be performed at the same time as revealing the results of a previous trick, or when shuffling cards prior to showing a trick.

Making Use of a Lack of Expectations One of the 'rules' of magic, not to repeat a trick twice is key to the re-framing of narratives in time as when audiences do not know what to expect, they do not know where to be looking, so are more able to be misdirected [127]. The flip side to this is that if the magician actually performs a set of actions once, then when they later only pretend to perform those actions, they have built up an expectation in the eye of the audience for what they are about to say [71].

Expectations are also key to recovery from a breakdown in a trick. If the magician can shift their internal narrative into a different trick without a break in their external narrative, the audience need never know that they are missing a trick [99](p34). The need for this lack of expectations is an important part of the magician's rule that a magician should not tell non-magicians the secrets of their tricks; however it is worth noting that a truly skilled magician can perform even the simplest tricks to a knowing audience and still surprise them.

People may have different beliefs about the same situation Many stage magicians have fought against the fact that a significant proportion of the public viewing their tricks will believe that they are being achieved by paranormal means. As shown by Hergovich [59], even when explicitly told that they are seeing a trick, some people will believe it is paranormal. This fact has also been exploited by mediums and other spiritualists [66].

Deceptive techniques have ethical implications Several magicians have actively campaigned against false mediums and spiritualists who they claim are using these techniques in order to prey upon bereaved people. Whilst to some extent this may be done due to professional interest as it helps maintain the suspension of disbelief that Binet describes as being vital to the entertainment of magic [13], it is clear that the hidden use of these techniques without framing them as tricks will in some situations be morally dubious.

3.4 Examples of Magical HCI Experiences

Magicboarding [96] is a way in which children can provide input to a design process by creating visual storyboards. It differs from traditional storyboarding in that a computer is used which accepts natural voice commands as to what to draw and then magically draws these ideas into a storyboard. This design removes a barrier of traditional storyboarding, which is that children lacking drawing skills find it hard to participate. The way in which this magic drawing works is by a remote connection to an artist who is able to hear the commands and draw them on the screen. The hiding of the person and framing the system as magic is useful here as it makes it seem less like an adult taking the children's ideas and developing them and more in fitting with the participatory design ethic this system is trying to foster.

Another project that employed deception was Anywhere-Everywhere-Somewhere [143], an interactive city tour by Willi Dorner & the University of Nottingham. In Anywhere, the participant is given a mobile phone to carry and at various points in the city experiences occur. At one point the participant finds a mysterious door with a piece of paper showing their own face on it and a video is displayed on their mobile phone showing what is behind the door. At another point, the viewer is told to navigate the city by following a trail of faces which occur in the environment, one of which is actually their own face, shown on a screen in a shop window. These two effects are achieved in a very different manner. The picture for the printed face is taken as part of the pre-experience briefing, and is printed out and stuck to the door by a person working on the project. This is a classic example of re-framing the experience in time so that preparation for the 'trick' is occurring whilst the participant thinks they are just doing some administrative tasks. The picture that appears on the shop window screen is actually taken automatically by the front facing camera on the device that the participant is carrying, at a point in the tour when a picture is shown on the screen with small details, which require the viewer to look closely in order to see the details. This gives the illusion of a picture which must have been taken up very close to the person, which clearly they have not seen occur. This is essentially hiding the setup for the trick in the actions of the user themselves; they do not know that they are being encouraged to be in the right place to take a photo. The hardware helps in this illusion, as the main, obvious camera on the phone is on the back, the user facing camera (designed for video-conferencing) is a tiny square at the corner of the display, which users are likely to be unaware of.

Anywhere also highlights the potential issues with using trickery without warning as one participant found parts of the experience seriously disturbing. Several changes were made to the framing of the experience during the run, from the initial framing which was simply as 'an interactive city tour' with no further information, which made the use of human performers and orchestrators and deception very surprising. This demonstrates a potential risk with the use of deception in settings where people are not expecting that trickery will occur; without that knowledge, they may find tricks unnerving or intimidating.

Magic Made Fun - a piece of software for the Nintendo DS handheld gaming system, allows a person to perform magic tricks, where the handheld device supposedly performs impossible feats, with the help of the performer. As an example, in the ‘Blank Card’ trick, the audience member is allowed to pick a card from a (physical) pack of cards. This is then placed face down next to the DS. The performer draws a rectangle on the touchscreen of the device. After a brief pause, the image of the back of a card appears on the screen in the middle of the rectangle, which then flips around to reveal a picture of a card. When the physical card is turned over, it is found to be the same card shown on screen. This trick uses a combination of trick equipment and hidden communication between the performer and the computer to produce the trick. This trick will be analysed in more detail in Section 3.6.

It is also important to note when considering technology and illusion, the tradition of stage magicians making use of cutting edge technology in their stage shows. For example Robert-Houdin famously used electromagnetism and other cutting edge technologies in his stage shows [111]. He also sometimes claimed to use other new technologies such as anaesthesia as a cover for effects that were actually performed by much more mundane means.

3.5 A Deception Framework for Describing Illusion

3.5.1 Factors to Consider

In a magical situation, what is happening is that someone is being fooled, having an illusion created for them, or being allowed to create their own interpretation of events. There are several key variables which can be different in a deception, these are described in the following paragraphs.

Who is fooling who? Tognazzini [128] describes the design of systems to create a consistent illusion which may not be necessarily directly related to the underlying system. In this case, it is the computer fooling a person.

There are also situations where a person may try to fool a computer. For example in gaming, people try to fool computer artificial intelligence in order to beat the game. People attempting to access computers often try to fool the computer into thinking that they are an authorised user (with typically the root or administrator user being the most desirable.)

More complex situations occur when computers are used in a collaborative or network setting. For example a person may try to fool a computer system, in order to indirectly fool another user who is also using that system. Alternatively, a computer system may be explicitly designed to allow the user to provide false information to others, as in the Swarm system described above [114].

What do they know? When someone is being deceived, their level of knowledge of the fooling is also an important factor.

In some situations the person or computer being fooled may not even know the person fooling them exists, such as in the magicboarding example above. In other situations, they may know the person exists, but not know that there is any capability for deception available to that person, or not know that the person is deceiving them. Finally, there is the situation where people are 'fooled', yet they know that they are being tricked – much of stage magic falls into this category – the magician provides a convincing illusion of a situation which is inexplicable to the audience, yet most audience members are confident that there is a mundane explanation underlying it [13].

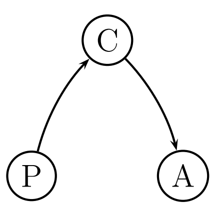
Why are they being fooled? An important question to ask when designing systems that create or allow illusion, or when exploring the possibilities for illusion is why the illusion is required, and what effect it has. This is particularly important when the person is not told that they are subject to an illusion, in which case care must be taken to consider the effects both on the performance and on the user themselves if the deception is discovered. If a revelation of the deception is planned, consideration must be made of the effect on the person being tricked, and whether there is a justification for this effect.

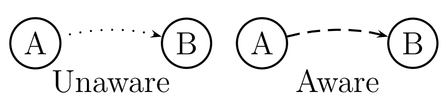
Similarly, when looking at analyses of existing illusions it is important to consider the motivation of the people creating the illusion. In most of the examples given here the primary motivation will be as part of a performance or other entertainment. However, clearly many uses of deception may be motivated by other factors and being clear about what these are in a particular situation is important. For example when people break into computer systems or attempt to gather user's bank details or passwords, the primary motivation is financial – this realisation has led to the relatively new computer security trend to apply economic and social concepts to the analysis of computer security threats, in contrast to previous strategies mainly inspired by military warfare [2].

3.5.2 Types of Deception

In the following sections, several different categories of deception involving either a person and a computer, or two people are described. This is followed by a description of the types of 3 party deceptions possible in a typical performance, where an audience, a performer and a computer system are involved in deception. These are comprised of combinations of 2 party deceptions. Examples are given, primarily from performance and gaming, with uses from either mainstream computing or performance use of other technology where an interactive performance example could not be found. A diagrammatic form to describe these deceptions is shown - this is based on the cooperative work diagrams introduced by Dix to describe interactions between multiple people and a computer [37]. Again, as described in Chapter 2, the computer is shown as an active participant in the performance, due to the potential complexity of the role which it takes.

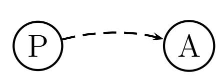
The situation is the same as the basic performance situation described in Chapter 2. Two people, Audience(**A**) and Performer(**P**), are connected via some kind of computer or technological mediation (**C**). So, in a typical situation, with no deception involved this communication would be shown as per the first diagram below, with solid arrows showing honest communications. Two types of deceptive communication are also shown by the different arrows below - a dotted arrow represents deception where the target is unaware of the deception, whereas a dashed arrow represents a deception where the target is aware that the deception is occurring.





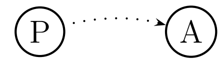
3.5.3 Two Person Deceptions

Knowing Deception (Suspension of Disbelief) Here the performer is deceiving the audience member who knows this deception is occurring. In performance and entertainment settings people may often deliberately suspend disbelief, for example in stage productions, events occurring on stage are

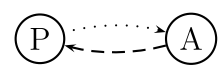


known not to be real, but audiences are willing to understand them as real within the context of the story.

Hidden Deception Sometimes a performer deceives an audience who does not know the deception is occurring. In magic performances hidden deceptions are often used, for example a person is often deceived into thinking that something is held by a performer in his hand, when it in fact is in his pocket or the other hand. Sometimes (as in most magic performance) the hidden deception is framed as part of a performance where the audience knows that a deception is occurring. In some situations lack of framing is not a problem, white lies are often used by performers, for example street performers often ride large, 5 or 6 foot tall unicycles, which are inevitably described as being 10 foot tall. However, deceiving the audience completely may not be the right thing to do as false mediums and similar scams demonstrate.



The Big Con This more complex situation demonstrates the interesting effects that may occur when several different types of deception are combined. The audience member thinks that they are deceiving the performer, however the performer knows this, and is also deceiving the audience member in return. David Maurer's 1940 book The Big Con [93] describes the theatrical methods used by con-men in the USA in the 1940s, which were perhaps a forerunner to the internet scams of today. These, whilst not being part of the mainstream theatre are worthy of description here as they required very similar skills to many performance works. Maurer describes how most con games involve a con man convincing the person being conned (the mark) that they are in fact taking part in a scheme to defraud a third person. Some of these were simple 'short cons' such as 3 card Monte, where a player has to guess which of 3 playing cards is a Queen. Whilst ostensibly the mark is playing against a single person, in practice one of the onlookers is a confederate, and says he will mark the queen somehow for example by bending the corner of the card. The person running the scam knows this however, and uses a sleight of hand technique to bend the card back and mark another card. There were also 'big con' games where entire fake gambling shops, telegraph offices and stockbrokers offices were created in order to convince



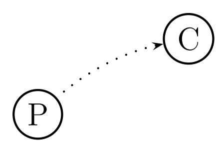
the mark to invest serious amounts of money in fraudulent enterprises. The con men used the fact of the mark's deception in order to protect themselves from problems once the mark lost their money - as the mark could hardly complain if their attempts to cheat the system failed to make them money.

This idea of deceiving a person by making them think they are actually taking part in some kind of dishonesty is also described by Randi [106] as a magic trick aimed at children, where the magician pretends to have been caught in the act of a trick by a child, whilst in reality, this is simply a setup for a real trick. In magic terminology this kind of trick is known as a sucker trick.

3.5.4 Deceiving a Computer

In this section, examples are taken from games, as this kind of single person deception does not on its own have an obvious performance use, except as an element of a computer mediated two person deception - which will be described in Section 3.5.6.

Deceiving a computer When the word deception is applied to technological systems here, it means that the system has a particular set of inputs, and the underlying assumptions of the system lead it to believe that a particular action is causing these inputs, whereas in reality the action being performed is different. In a very simple case, if a person logs in with a stolen username and password, they are 'fooling' the computer into believing that it is being accessed by the legitimate user. One example of this from computer gaming involves ways in which people beat the artificial intelligence built into the system. For example in the early computer game Gauntlet [139] enemies would always move towards the player in a straight line. This meant that by moving behind a wall, the player could trick the enemy into acting as if it had no way to get to the player (Figure 3.1).



Deceiving a Knowing Computer In terms of knowing deception - if the computer is aware of the deception - it simply means that the computer has some way of detect-

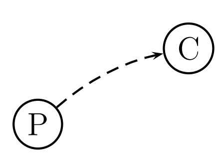


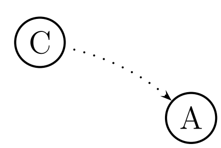
Figure 3.1: A player moves round a wall to stop a ghost attacking

ing or expecting the deception, for example a system may be designed so that it will take a particular set of inputs, and detect inputs which are outside that range.

For example, some games are designed with auto-fire detection, to stop people using hardware or software tools that allow them to simulate multiple very fast button presses very easily[1]. A system designer may also deliberately choose not to limit 'impossible' inputs, which are not possible to create using their expected range of legitimate inputs, even if the computer can detect them. In these cases, the computer system could be said to be suspending disbelief (this does not necessarily require any high level artificial intelligence on the part of the system - more accurately, the designer is choosing to make the system seem to suspend disbelief). This may be particularly useful as part of a design strategy where it is hoped people will interact with the system in ways not envisaged by the designers [116].

3.5.5 Deceived by a Computer

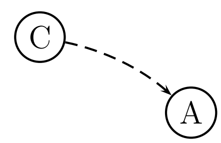
Deceptive Computer Computer systems may also be described as 'deceiving' users, although in many cases it is essentially the entire computer system, including its designer who are achieving this deception. Again no intention or strong artificial intelligence is implied of the computer itself. A computer deceiving a person is a system which (by design or otherwise), provides outputs to a user which are designed to create a belief in the user that a certain state of affairs



[1] e.g..http://www.emulinker.org/index.php?page=Features

is true, when that is not the case. Game design often uses deceptive elements - for example in many games players will see a seemingly empty room, and only once they get into the room will monsters appear, having been hidden behind walls or been otherwise out of sight (the Zelda series of games by Nintendo have particularly inventive ways of performing this empty room deception, such as a creature breaking through a wall, or rising out of a pool of water [155]).

Suspending Disbelief of a Computer Tognazzini [128] argues that all computer programs essentially aim to deceive the user into following their model of a situation. In this case he is not talking about deceiving in the strong sense; users know when they interact with 'pages' on the internet, or 'slides' in a PowerPoint that the computer is not really holding a stack of paper beneath the screen. This is more about encouraging people to suspend their disbelief, and creating a convincing fantasy of the operation of the program which allows it to work effectively. Games are perhaps the most explicit examples of this type of suspension of disbelief deception, they encourage people to believe in entire fictional worlds, computer generated characters, and other clearly impossible things.

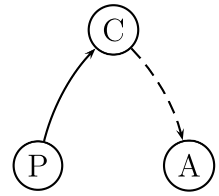


3.5.6 Computer Mediated Deceptions

The deception patterns described here are combinations of the two party deceptions above, placed in a performance situation, with a performer, and audience, and a computer system involved. Each is presented with a diagram of the deception, a description of the deception, and a description of types of strategies which this may support in a performance setting.

This section considers combinations of deceptions which flow from performer to audience via the computer. In real situations, a combination of these strategies will be used, potentially in tandem with direct person to person communications and deceptions.

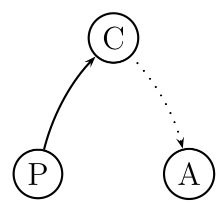
All 8 possible combinations of computer mediated deception are described here, although as explained at the end of this section, 1 combination is not possible in reality, so will not be discussed in later sections.



Story Telling In this situation a performer (P) is controlling a piece of technology in order to deceive an audience member (A). In this situation the person knows that they are being deceived. Typically in a performance situation, they will be suspending their disbelief.

As an example, in the 19th Century automatons were used by magicians. These were machines deliberately designed to fool the audience into thinking something amazing was happening. Probably the most famous of these is Robert-Houdin's 'Orange Tree' illusion, where a small orange tree appeared to grow first flowers and then oranges within the space of minutes. These automatons were either controlled by the performer setting off clockwork mechanisms or by offstage assistants using levers and wires to control them. Magicians were very open about the trickery involved, so people were aware of being fooled, although they did not know about the exact technology being used. Similarly in theatre, illusion involving technology is often used, such as blank shots from guns, disappearances through trapdoors, sound effects etc.

This type of deception is useful for supporting activities such as story telling and theatre, where we are not wishing to actually deceive a person, rather to create an impression that fits into the story being told.



Magic Trick Here, a person is using technology in order to fool another person who does not know that they are being deceived.

The facility in Desert Rain to 'nudge' players and have a hidden effect on their progress in the virtual world is an example of a performative use of technology deliberately designed for deception. As Blast Theory describe [78] care has to be taken when designing this kind of deceptive control into interactive games, in order not to remove the feeling of agency for the player, and to avoid letting the player know this control is occurring. They completely hid the performer, so that the player had no idea that they even existed.

The frame of a magic performance is an open deception - where people know that they are seeing tricks. However, within that performance magicians often deceive the audience without their knowledge, for example in the use of gimmicks; pieces of technology designed either to look like normal items, or to be hidden, which are specially designed to perform a trick with. One computerised version of this kind of deception, is the MOAOCT (Mother of all Online Card Tricks) by Jason Litchfield [149]. This uses a computer to supposedly mind-read a selected card, with all visible interaction with the computer being by the audience. In fact, what is actually occurring, is that an assistant either off stage listening in, or nearby and using a small portable device such as a mobile phone, uses a different web page to simply set which card is shown.

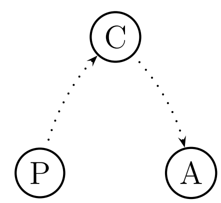
Entertainment magicians have campaigned against the use of technology by spiritualists in order to produce 'supernatural' experiences, which highlights the importance of framing this kind of deception and being careful when using it. However many uses of deceptive technology do not have ethical implications, such as the use of juggling chainsaws by circus performers - these typically look like normal chainsaws (and sometimes are actually real chainsaws), but have toothless chains, meaning they will not cause major injury in the event of a mistake. Whilst the audience may not know that the chainsaws are modified, they do trust that the juggler isn't really in serious danger, and suspend their disbelief of this - an important part of performances involving danger is in tempering an audience perception of danger by creating trust in the professionalism of the performer so they believe that the performer is skilled enough to not really hurt themselves. In this way, a deception of the audience about a particular device is transformed by framing into a knowing deception about the level of danger involved.

In general computer use, some systems such as instant messaging clients actively support deception - for example by providing an 'appear offline' status which allows users to be online, view the status and make outgoing connections to others, whilst those on their friends list see their status as being offline. Systems may be used for deception without actively participating in the deception, for example it is easy to tell lies via email systems without having to use any special deception function.

This kind of deception supports strategies where we wish the person to actually believe something has happened which has not, such as the Desert Rain example, where we wish people to believe they moved themselves smoothly to a destination, when in fact they required some level of help from the performer. In that case, the deception is simply for orchestration of the performance. However, we may wish to perform using this kind of deception to create effects which seem impossible or seem to be far more virtuosic than they really are.

Whilst it is sometimes possible to create this kind of deception using standard systems, in many cases use of this kind of deception may rely on having the ability to modify the system being used in order to allow it to support the deception. This may limit the use of this kind of deception in cases where the system being used is not customised.

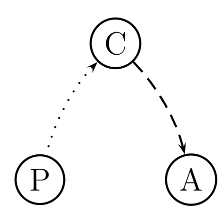
Beat the System Another way to deceive a person is by fooling a technological system and using this deception to fool in turn the person they are connected to via this system. This differs from the deceptions above in that the technology is not an active part of the deception - instead it is being fooled itself.



One example of this is the use of tricks in multiplayer games to exploit ways in which the system works. For example, in some online multiplayer games it can be advantageous to cause a network delay so that the other player is temporarily unable to see what you are doing. This kind of deception is also key to many non-performance computer deceptions, such as unwanted intrusion into systems, where a computer system is tricked in order to allow the attacker to impersonate a legitimate user.

This type of deception is useful for similar performance strategies to the Magic Trick deception above. It has one big advantage which is that it does not require modification of the technical system itself. However, it is limited depending on whether weaknesses can be identified in the technical system.

Stupid Computer In this situation, the computer is fooled by the performer and is passing on false information based on this, yet the audience member is able to understand that this information is false.

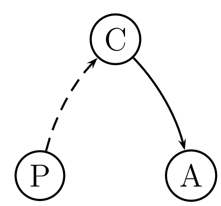


In a performance setting this may be a failure to deceive the person, which could be a problem. However, it also may be possible to use this deception of the computer as an integral part of the performance; one example is the use of multiple people hitting pads with their hands in order to complete levels in the computer dance game Dance Dance Revolution [138], which is done for the entertainment of those watching. Whilst the computer thinks that the dance pads are being hit by a skilled dancer, to the audience it is clear what is going on, and comedy ensues.

In general computing, this kind of deception occurs regularly. For example in the event of a computer being hacked to show a fake internet banking website, factors not knowable by the computer may allow a person to detect that the site is false, for example poor English grammar, or logical errors in the behaviour of the site.

As a strategy for performance, this may create comic effects, as in the Dance Dance Revolution with Hands video referenced above. It may also be useful for situations similar to that in the Story Telling example, where the performer wishes to put across the idea of doing something without actually doing it, as in theatre, and the audience suspend their disbelief.

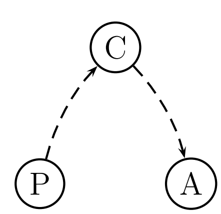
Caught by a Computer If a person attempts to deceive a computer, yet the computer knows this deception is occurring, it may pass on this information - to warn others of the deception occurring.



The auto-fire detection for games described above may be used to warn people that they are playing against someone cheating. Similarly, email spam filtering or web browsers which detect malicious or hacked web sites are examples of this kind of failed deception.

In terms of a performance strategy, this deception may be useful in performance settings where an audience member is interacting with a computer. An example of this will be shown in the Cup Game (Chapter 5). Using systems that have functions to catch trickery may also be useful as a way of demonstrating honesty, in order to make a trick more impressive. This is similar to the way to the way escapologists often chose sailors or other experts to check the knots that they were tied up with [65].

Computer Suspending Disbelief This is a computer that is being deceived and knows it - yet chooses to pass this on.

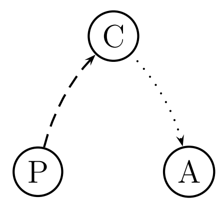


In a performance setting this may be a computer that is designed to take a certain set of inputs, but deliberately designed not to detect when an input is outside that range, or to interpret out of range inputs as something within their expected range.

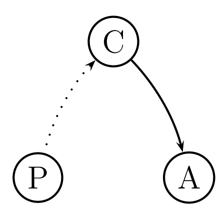
As a strategy, building in support for this kind of deception may be useful in order to allow users to play with the inputs they create, for example by allowing multiple users to pretend to be one user for the creation of comedy effects, even if the system can detect this. An example of this kind of trickery will be shown in the juggling tracking system (Chapter 4).

In some ways, this suspension of disbelief by the computer system is actually similar to designing in computer trickery as in the Magic Trick deceptions. The difference is that the computer is not explicitly designed to take a particular envisaged input, rather it is designed not to disallow inputs outside the expected range of inputs.

Sneaky Computer This is similar to the Computer Suspending Disbelief example above, except that in passing on the deception the computer does not make it obvious to the audience member that the trickery is going on. This may be the case if the audience member cannot see what the performer is doing, for example if the audience member can only see the output from the computer and not what the performer is doing to make it work.



Not Really a Chain of Deception In this situation a computer is tricked without knowing it, but is not passing on the deception to A. The only way this situation can occur is if the communication between the computer and audience does not relate to the communication between performer and computer; this is not really a chain of communication, more two separate deceptions. As such, this type of deception is not discussed further in the following chapters.



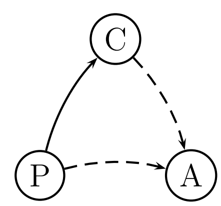
3.6 An Example of a Computer Mediated Deception

As an example, we will break down the deceptions involved in the Blank Card Trick from the Nintendo DS Magic Made Fun, as described in Section 3.4. This will demonstrate how several deceptions may be combined in order to make a single deception. The trick itself is a knowing deception, where the audience member knows that they are being tricked. However, this trick is made up of several deceptions which the audience member is unaware of. All the vital parts of the trick happen in full view of the audience member.

What appears to happen in this trick is:

1. The audience member picks a card and puts it face down on the table.
2. The performer draws a rectangle on the screen of the Nintendo DS.
3. The DS pauses, and shows some animation, and then shows the card, which is seen to be the same as the chosen card.

This is framed as a magic trick where people are being knowingly deceived by the performer, both directly, and via the mediation of technology. A deception diagram for the whole trick is shown on the right.



The deceptions which actually happen in each section are described below.

3.6.1 The Audience Member Picks the Card

In this part, what the audience member thinks they are doing is simply picking a card and putting it on the table. They do not have to look at the card at this point for the trick to work and the card deck can be inspected and shuffled as much as they like.

Once the card has been chosen and put on the table, the first part of the trick is worked. In this, the performer ascertains which card is face down on the table. This is made easy by the fact that the cards used are a marked deck, with the suit and number written subtly on the back of the deck (Figure 3.2).

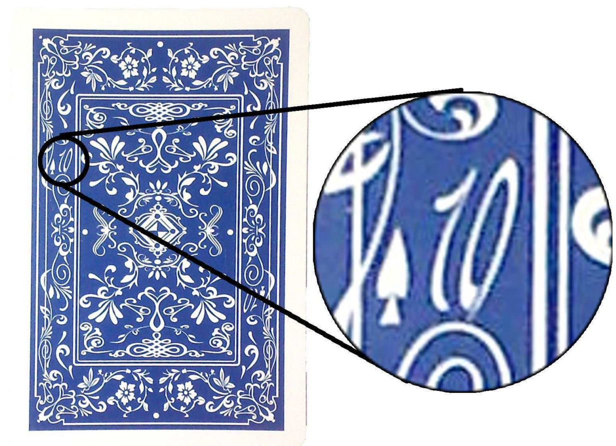
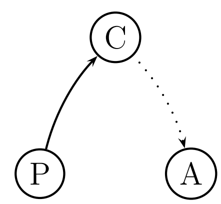


Figure 3.2: Marked deck of cards

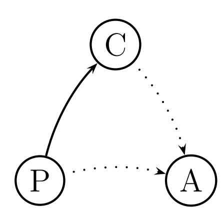
In this part of the trick, the performer is using the card technology in order to fool the person that the card itself is unknown to the performer. This is the Magic Trick deception shown to the right , even though the cards themselves are physically being handled by the audience member, the deceptive technology is being used by the performer.



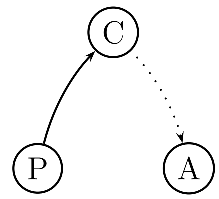
3.6.2 The Performer Draws the Rectangle

When the performer draws the rectangle on the screen, they act as if they are just naturally drawing four lines. However, what the performer is actually doing, is drawing four lines in a particular order and direction. The handheld console interprets the order of line drawing using a code built into the game, in order to say which card it is.

This section is a combination of the performer deceiving the audience member into thinking that they are just drawing lines in any order and that this has no bearing on the trick itself (which is in itself a hard trick to learn), and also of communicating via the console in order to create the deception.



3.6.3 The Computer Guesses the Card



At this point, the computer appears to guess the card. In reality, it knows which card it is already. The computer has some animation in this section so that it seems to be doing something, after which it shows the image of the card.

Here, the performer is simply using the computer to perform the trick, another Magic Trick deception.

3.7 Conclusion - What is It For?

In this chapter, a framework for analysing and describing computer mediated deceptions has been demonstrated, along with suggested design strategies which each type of design may support, and an example of a magic trick being broken down into its constituent deceptions.

In the next chapters, three designs will be presented which have been informed by this framework, and which demonstrate its usefulness in three very different performance situations.

The first, the Juggling Tracker, was built early on in the development of the framework, as such, the process of performing with this system inspired much of the framework development. The juggling tracker work demonstrates how deceptive techniques may emerge as useful techniques in a general performance situation.

The second design, the Cup Game, was developed when the framework was fully formed. It is an experiment in designing using the deception framework. It performs both as a demonstration of the use of the framework in a design process, and as an exploration of the issues involved in framing the use of deception.

In the third project, Rock, deceptions are explored in the context of an art installation. This provides insight in particular into how deceptions and suspension of disbelief work in situations where the roles of performer and audience are less well defined, and may change at any time.

Chapter 4

Eye-Balls: A Juggling Tracker

Eye-Balls is a system for augmenting juggling performance. It uses a video camera and image processing software in order to track a person and the multiple objects that they are juggling. It then creates output which responds to the person's actions. The output can be in the form of projections and audio, and it is also capable of controlling MIDI controllable devices such as lighting effects and automated stage props. This means that the juggler's actions can directly control both their direct performance, i.e. the juggling itself, and also the other audio/visual components of the performance. This is in contrast to the majority of methods of accompanying a performance, such as the use of pre-recorded music or video, in that the juggler has control over the accompaniment rather than having to perform their actions in time to the accompaniment. In some ways performance in Eye-Balls is a similar experience to performing with a live accompanist (albeit a slightly dimwitted one that requires a lot of teaching before it is ready to perform!) The juggler is able to monitor the system's responses both visually and to hear the audio which means that they are able to respond in turn to how the system is responding to them, creating a feedback loop between them and the system.

In this chapter the development and use of the system will be described. This will demonstrate the issues involved in concurrently controlling the system and also performing a juggling act, and the techniques adopted to make this work. These take account of the differing perceptions of audience and performer and demonstrate how deception may be an effective strategy when working within the unique constraints of performance interfaces like this, where there is no choice but to interact with the computer in full view of the audience.

4.1 Why?

Juggling performers typically fall somewhere between two extremes in terms of accompaniments. One extreme times everything perfectly and repeats a show exactly move for move. This type of performer can use carefully crafted accompaniments and equipment such as juggling props which change colour in a manner pre-sequenced to be exactly in time with music, or pre-recorded video projections. At the other extreme, typically for more comedic acts, performers will do a performance with a lot of talking to the audience, comedy patter and acting etc. This kind of performer may either not use accompaniments or will use music and juggle in time to it, but without pre-planning their performance in the same detail. This is because when you are interacting with the audience and improvising parts of your act, it is hard to calculate in advance exactly what time you will perform particular actions.

By making the technological parts of the performance responsive, Eye-Balls opens up the ability to perform in a less pre-planned manner but whilst still having the ability to have accompaniments which are a close fit to the juggling. For some kinds of accompaniment it can make things that are extremely hard to do significantly easier, such as juggling very hard tricks exactly in time with a musical track; this potentially extends the range of juggling tricks available to a performer to allow ones that take 100% of their concentration, which they could not previously do in time with their accompaniments.

In addition to this, Eye-Balls opens up completely new areas of performance where the performer's interactions with the system are actually part of the performance. For example in one performance in Eye-Balls, the performer played a game of Space Invaders on screen, with the firing at aliens being controlled by ball throws, and the movement of the spaceship being controlled by movement of the juggler's body. This meant that the juggler had to improvise his act around the alien spaceships coming down towards him, which were in turn responding to the juggling he was performing. A major part of the performance in this act was the display of the juggler's skill in performing stylish juggling moves at the same time as moving to avoid being hit by and to shoot at aliens.

This chapter provides an example of the way in which visibility modification, and also the use of misdirection and other strategies may occur in a live performance situation. The system has been used for several live performances during

Date	Venue	Audience Size
Jul 2006	Initial technology workshop	6
Jul 2006	Testing with a professional juggler	1
Oct 2006	Live demos at GameCity festival	50
Oct 2006	Further testing with local juggling club	10
Nov 2006	First full public performance - at Twaddle Cabaret, Derby, UK	100
Feb 2007	Testing with another professional juggler	1
Mar 2007	HitlabNZ Symposium, Christchurch, New Zealand	100
May 2007	Performance at Design Camp 8 - Wanganui, New Zealand	500
Jun 2007	Small group performance & individual demonstrations at Creativity & Cognition, Washington DC,USA	20 at performance, 100 at demos
Sep 2007	Performance & presentation, (re)Actor2 Conference, Leeds, UK	20

Table 4.1: The Juggling Tracker in Public

development (see Table 4.1 for a list of selected performances). These performances have helped to develop insight into how the system works in a real world situation.

4.2 Related Performance Work

Several performers have combined juggling and electronic technology in the past. Firstly, several companies have made juggling props such as balls and clubs which contain built in lights, allowing them to be juggled in darkness, creating beautiful patterns due to persistence of vision creating the effect of glowing trails in the air. More recently, programmable props have become available which have multiple coloured LEDs in, and allow the user to program a set of colour changes into the prop, which are then replayed in time with a musical accompaniment (several high profile performers have used these, see http://www.globall.com for videos). These props still lack the element of interactivity and the performer is forced to perform to a fixed accompaniment.

Several performers have created systems which add an interactive element to juggling. These mostly use some form of body attached hardware in order to trigger sounds. Various different types of hardware were used to achieve this, such as using arm flex sensors to sense tempo and play tunes in time with juggling [145], using muscle nerve sensing to create audio based on the juggler's movements [132], bouncing juggling props off arm mounted sound pads [158] and using juggling clubs with special infrared reflective markers, tracked with industrial high speed cameras [17]. Another related project uses 'poi', which are small heavy objects on the end of string, swung round the body in elaborate circular patterns. These contain accelerometers, and are used to create audio and projected visuals based on the spinning of the poi [117].

L'Universe, a collaboration between the Flying Karamazov Brothers and MIT [110] was probably the most technologically elaborate juggling performance ever. This used a combination of sonar tracking, accelerometers attached to the wrists of performers and sonar generating hats, along with wearable computers, to drive a 'stage instrument', which played music controlled both by the juggling and the performers' movements on stage. They also used wirelessly controlled juggling clubs, which could illuminate in multiple colours in order to provide visual displays controlled by the juggling pattern. This system involved a large and varied set of technological equipment, and took several hours to set up.

Eye-Balls is designed to improve on these past systems in several ways. One key way is by being fast and simple to setup, and by only using standard equipment. This is in order to enable it to be used in real world situations by typical juggling performers and groups, who usually do not have the funding, or expertise to work with non-standard equipment, but usually do have access to a laptop and a video camera or web camera. It is also designed to have an extremely flexible scripting mechanism, in order to allow for more complex performances without the necessity of external assistants.

4.3 Technology

The juggling tracker uses a video camera connected to a computer to track the positions of the head and hands of someone juggling, and the objects which they are juggling. This is then fed into a script based output system, which creates a response to the juggling, which can be output in the form of video projections and audio. This section discusses the technical details of the system.

4.3.1 The Tracking Algorithm

Eye-Balls uses a standard video camera as input, either a DV camera or a web cam. The DV camera is used in large scale stage settings, where the juggler may be further away from the camera, and the extra quality of the camera and higher sensitivity of the larger lens and sensor are required. The camera is attached to a laptop running a custom DirectShow video processing filter which processes the input.

The processing filter uses a new particle filter based method developed during the project to track the balls and the hands and head of the juggler. This method is innovative in two ways. Firstly, it uses a simple partitioning method in order to track all visible balls using a single particle filter. Secondly, it uses information from the hand and head trackers in order to aid the ball tracking, and vice versa. The algorithms are described in terms of their practical implementation below. The mathematical details of this method were published at BMVC 2006 [89] (included as Appendix A).

Particle Filter A particle filter (e.g.the CONDENSATION Tracker [68]) uses a set of 'particles' each of which stores a hypothesis as to the position of the object being tracked. At each time step, each particle is scored against the video frame as to how well the particle fits the real data. The particle set is replaced by a new set, constructed by weighted random sampling from the original particle set, weighted by the particle scores. This new particle set is then updated using a motion model, for example each particle may store a velocity and may assume that the velocity is constant, meaning that each particle is moved by the velocity stored in the particle in the update step. The update step will also add a slight amount of random perturbation to this movement (and to the calculated velocity in this case). This randomness creates a local search around the expected position.

The multiple particles give it an ability to account for multiple hypotheses which is useful in that it allows mistakes to be recovered from, so if most particles of a tracker miss an object for a frame, if even a single particle tracks correctly it may be re-tracked in the next frame.

The particle set is in effect an estimated distribution of the probability of each of the object positions at any time. This is typically used to give a single output for the best estimate of object position, for example by taking a mean position over all particles. Because multiple balls are being tracked in Eye-Balls, a custom partitioning method is used in order to give estimates for all ball positions at once, and to keep this distribution multi-modal.

Partitioning The tracking of the balls uses a single particle filter with a particle set containing particles which each contain the position of a single ball. When multiple balls are visible this particle set will be a multi-modal distribution with peaks at each ball position. In a simple particle filter this distribution would converge over time and lock onto whichever ball was tracking most strongly. To avoid this, before the weighted sampling is taken over the particles the particles are clustered in order to detect the multiple modes of the distribution and each cluster is sampled individually. This partitioning is done using a simple distance criterion, which is that two particles are in the same cluster if the two potential ball positions they are describing overlap. The mean position of each cluster is output as a ball position. Clusters with a very low score are assumed not to currently point at a ball, so are not output as ball positions. This allows for the tracking of multiple changing numbers of balls without having to explicitly tell the system how many are being tracked.

A similar partitioning algorithm is used for the two arms; in this case, the particles are split into two clusters, one on each side of the performer's head, making the assumption that the performer is facing forwards.

Tracker Information Sharing As well as tracking the balls the arms and head are also tracked. These are tracked both because they provide interesting information as to when balls are caught and the position of the juggler, and also because they can be used to improve the ball tracking.

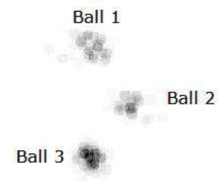
The ball tracking is most challenging at points when the ball is caught. At these points, the hand tracking is used to help the ball tracking work more efficiently.

In turn, the ball tracking can also be used to improve the tracking of the hands. This means that the overall tracking is improved significantly compared to the use of standalone particle filters for hands and balls.

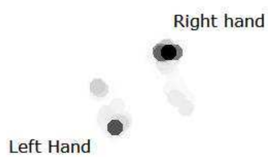
The interaction between the ball, arm and head filters is done using a novel algorithm to represent tracker interaction. Firstly, for each particle set, after weighted sampling and motion update, a 2 dimensional matrix is generated representing a quantised version of the 2d space in which the tracking is being done, i.e. the image plane. For each particle the parts of this matrix which are at positions that are covered by the particle's object are incremented by an amount relative to the score of the particle. For example, for balls, a circle of matrix bins beneath the particle's estimate of the ball position and size are incremented. This is fast to do, as the matrix may be treated as a bitmap, making it basically a simple image processing operation. Once this is done, the matrix holds a representation of the probability of an object being in each position. Because this is taken after the motion update, it represents the predicted positions in the next frame. This means that it can then be used as input to the scoring functions for the other particle filters in the next frame. Because objects are typically only affected by other tracked objects when they are close to each other, the extra work in the scoring functions is minimal as only a small area of the interaction matrix has to be checked for each particle. Figure 4.1 shows the matrices for a particular situation.

This use of interaction matrices allows for better tracking, particularly at points when a ball is caught where ball tracking may be less reliable as the motion is less predictable. It also allows the use of a very simple arm model which when used with no balls being juggled typically drifts off onto another skin coloured body part and loses tracking relatively quickly. When balls are juggled and the interaction matrix is used, the assumption that when balls are caught the hand will be in the same place as a ball is used to effectively re-initialise the arm tracking. The interaction matrix is also a fundamental part of the arm tracker as the particles representing the two arms are partitioned by using the head position from the head tracker's interaction matrix. The head tracker actually has a simplified interaction matrix for this purpose that is basically 1 dimensional, being the probability of the head being to the right of a position, each head particle adds to only the matrix bins to the left of the head. Interaction matrices allow for enhanced tracking without taking significant extra time.

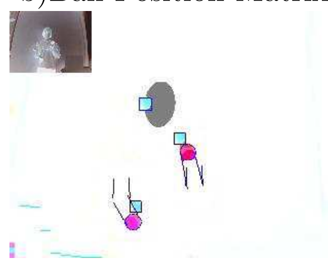
a)Input Frame



b)Ball Position Matrix



c)Hand Position Matrix



d)Output - Arm and Ball tracking

Figure 4.1: The tracker state for a single frame

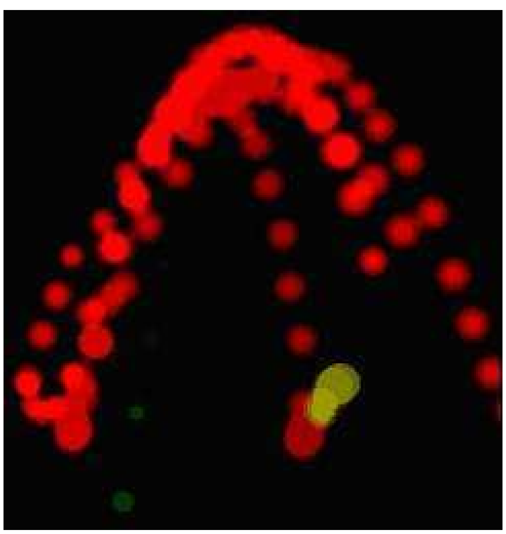
4.3.2 Script System

The data from the tracking algorithm is fed into a script system which controls what response is made to the input. A script is called at each video frame and is given the positions of the balls and the juggler. This script system runs a JavaScript interpreter, with several custom objects defined which control creation of output and handle pattern recognition tasks. An example script is shown in Figure 4.2. Each object in the script processor is described here.

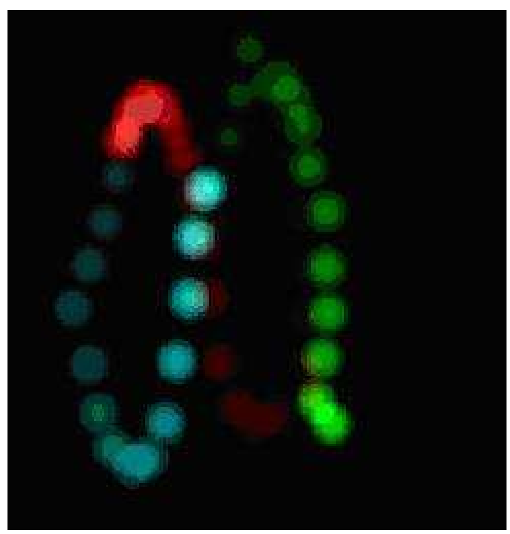
Engine Object The engine object handles the core tasks of the script engine such as registering which function(s) will be called each time a new set of data arrives. It allows the script to include other scripts which allows for a reuse of existing code and modularity in the design of scripts.

Display Object The display object handles visual outputs. These are created in two ways, firstly there are several provided input video streams including the raw input video, and display of coloured circles at the ball positions. Secondly, a sprite based system for adding animated objects is used in order to allow the addition of elements such as text and images to the system. The various outputs are mixed together using a 'processing tree' which allows multiple inputs to be combined and altered using filters such as video mixers, motion blurs, matrix transforms, and

```
if( !this.initialised)
{
  // initialise the display
  // draw circles where the balls are
  this.balls=display.Add("BALL_CIRCLE");
  // add motion blur
  this.blurrer=display.Add("MOTION_BLUR",
    this.balls,225,225);
  // output blurred ball display
  display.Add("output",this.blurrer);
  this.initialised=true;
}
if( ballCount==1)
{
  // only one ball in display:
  // blur length long so performer
  // can paint with the ball
  display.SetParams(this.blurrer,255,255);
}else
{
  // otherwise: blur length short
  display.SetParams(this.blurrer,225,225);
}
```



a)Painting with 1 ball



b)Juggling 3 balls

Figure 4.2: An example script and outputs

finally to be output to either the main projected display or a monitoring display seen by the performer.

Audio Object The audio object creates audio outputs. These are either waveform audio or midi messages. Audio processing is embedded into the system in order to allow waveforms to be sped up, slowed down, re-pitched and mixed together. This allows the system to use waveform audio but remain responsive to on-screen actions. The support for midi messages also allows for interfacing with stage automation and lighting control systems.

Pattern Object The JavaScript scripts themselves can perform simple pattern recognition, such as detecting how many balls are visible, or when balls are moved into certain regions of the camera's view. The pattern recognition object allows for the detection of particular juggling patterns. This is used because this detection was not fast enough in pure JavaScript. The pattern object uses internally a very simple k-nearest neighbour classifier [31] in order to detect which pattern is being juggled based on a set of training data.

4.4 The Juggling Tracker Design Process

4.4.1 Designing with Jugglers

In order to design a system with this wide a range of flexibility and to envision the possibilities of such a system, it was vital to have input from the wider juggling and circus performance community. Because of the need for external input a collaborative design process was used for this system. Several workshops were organised with local juggling groups and some time was also spent with two professional performers. The people in these sessions covered a range of experience levels from world class performers to hobbyists.

The workshops provided valuable inspiration in the early design stages as to what the sensing and pattern recognition should be able to do, and in terms of designing the output system. However, the point at which they were most useful was later in the design process, when the system was relatively stable and actual performances were being created within the system. Talking to performers who use existing non-interactive technology, such as programmable colour changing juggling balls, video projections and music was very helpful. In particular the discussion of issues we had about using technology during performance addressed practical problems such as setup time and synchronisation, and more performance related ones, such as the division of audience attention between the performer and the clever visual elements they are using. They also expressed concern about using technology because it is there, rather than for any artistic purpose.

Talking to Joel Salom, an Australian professional juggler who uses a home-made set of sensing pads attached to his arms and hands was particularly useful. His show has been developed over several years and is probably the only professional circus show using computer technology purely controlled by the performer in this way. His sensing pads trigger sound and lighting effects when balls are bounced off them, allowing him to create music entirely triggered by his juggling. A major issue that was brought up in this conversation is that of audience understanding. He found that at times, despite him explaining the system in advance, audience members would sometimes compliment him afterwards on how well he juggled to the rhythm, or would not understand the link between the pads and the sound effects. In an early public performance with the Eye-Balls system at a local cabaret I also experienced this confusion; several of the audience members I spoke to said that they were initially confused as to what the link between the pictures and the juggling was, with a various opinions being expressed as to which point in the performance they understood that there was a link. In future performances, I worked hard to start with a very simple example (as shown in the Bicycle performance in Section 4.5) in which it is obvious that the projections are responding to ball movement, which reduces this confusion.

4.4.2 Design Constraints

The system was deliberately designed without external control interfaces, so everything in the system is controlled by the juggler's movements. This decision was made in order to force users who wished to perform scene changes or otherwise alter the way in which the system is working at any point to implement the control of these through the juggling. This was done in order to avoid the discrete patch based methodology of systems such as Max/MSP, which lends itself to sudden jumps in the output as patches are changed, and to encourage script writers to experiment with more gradual or subtle scene changes.

From a hardware point of view, there are many custom hardware solutions, such as accelerometers, ultrasound positioning systems, 3d cameras, infrared beacons etc. which have been suggested as alternatives to the vision led approach used in Eye-Balls. The vision approach however has a major advantage in that it requires only very standard hardware, which most potential users already have access to. In practice it is also extremely reliable, and can be set up in most settings in well under 5 minutes.

Glow Balls such as Aerotech Globalls [136] are very popular amongst jugglers. These are balls that are lit up from inside with coloured LEDs. These allow juggling to be performed entirely in the dark, and create an effect of trails of light in front of the juggler as the balls move through the air. Due to the popularity of these props, a simplified version of the tracking algorithm was also created, which tracked only these glowing balls, in complete darkness. This simplified tracker is very reliable as there are no background distractions in the video image, and the brightly glowing balls shine out through the fingers when caught, so remain visible at all times.

4.5 An Example Performance

As it is hard to describe purely in words what the system does, an example performance is presented here as an illustrated script. The illustrations show the juggler's actions, and the video output created by the system at that point in the script.

Joe: *Hi, My name's Joe and I'm a juggler*

(I move a ball in the shape of the word Joe, which causes it to be written on the screen behind me.)

J: *Oooh isn't that clever!*

(jumps into victory pose, arms raised. The screen changes to a picture of a bike.)

J: *I have a dream. No, not that one. This one is a bit more odd…*

(At this point, I talk all about my bike, and how much I love it and how great it is)

J: *Anyway, I had a dream about my bicycle, and that it was juggling.*

(I start juggling)

J: *It was performing intricate patterns.*

(Balls are shown circling above the bike)

J: *And I woke up, thinking, how on earth can a bicycle juggle?*

(I yawn, arms above head, triggering the screen to change again. This time a picture of an engine is shown. I start juggling in a pistons pattern, where the balls move up and down in columns, like pistons in an engine.)

J: *I know how a car would juggle.*

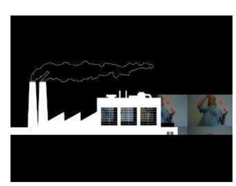
(The engine pistons move in time to the juggling, and engine noises are made.)

J: *High Gear*

(I do a double speed version of the pistons trick, and the on screen engine speeds up. This hard looking trick is followed by another victory pose, and the display changes to show a factory)

J: *This gave me another idea, what if jugglers were mass produced, and came out of a factory.*

(Juggling in a 'machine' pattern, as multiple images of me juggling come out of the factory)

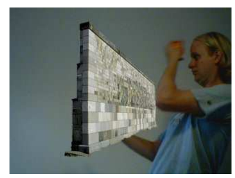


J: *Like a tennis match*
(Juggling in a 'tennis' pattern, where one ball goes high up from side to side over a net made by the other two balls. On screen, a tennis audience turns their heads from side to side to follow the high ball. Tennis sounds come out of the speakers.)


.....Various other dubiously stretched analogies follow, until...

J: *Then, I was banging my head against the wall*

(Banging head against onscreen wall, with comedy crashing noises.)

J: *and then suddenly it came to me. I knew how to juggle like a bicycle. But unfortunately it required the*
(throws hands up in horror)

J: *Hardest Juggling Trick in the World*
(The finale is a big and obviously difficult juggling trick that ends up with me in a position that looks like riding a bicycle, whilst still juggling.)

4.6 Visibility in the Juggling Tracker

In a typical setup of the juggling tracker, the performer is in front of the audience. The camera is in front of the performer, pointing towards them, as is a laptop running the tracker software. Speakers are somewhere on either side and a projection screen with the video outputs from the tracker on is either above, behind, or to one side of the performer.

In most venues, there is no way to physically hide any of the technology, and clearly it is not desired to hide the juggler from the view of the audience. In purely physical terms, the audience and performer are potentially able to see almost everything that is occurring in the performance, and everything that the juggler is doing is usually visible to the computer. The simplest analysis of the situation would be that basically everyone can see everything.

However, if we analyse how they are seeing things in more detail there are many significant differences which we can exploit for performance purposes. This section describes these differences and strategies for exploiting them which emerged during use of the juggling tracker.

4.6.1 Who can see what?

Computer Outputs Eye-Balls has three different outputs.

Audio: Audio is output from the main speakers, and is audible to both the juggler and the audience.

Video: Video is output on the projection screen. This video is also visible on the laptop screen.

Monitor Video: The performer has a separate monitoring screen on the camera, which only they can see, showing the unmodified camera view.

Performer Actions The performer, being human, has a wide range of outputs. However, we can roughly classify the outputs likely to be of interest in a performance situation.

Stage Movement: Moving around the stage.

Juggling: Throwing balls around in the air.

Body Actions: Non-juggling body movements, such as waving or stamping a foot. Some of these are visible to the computer, and all are visible to the audience.

Talking: Sounds made by the performer will be audible to the audience. Eye-Balls has no sound inputs, so will be unable to detect these sounds.

Audience Actions Audience outputs, whilst not detectable by the Eye-Balls system, are still vital to a performance in the system so must be considered when designing a performance.

Body Actions Such as clapping, sitting up, looking at a particular spot on stage. This type of cue is vital to performers to see how their act is going on. These will typically be visible to the performer.

Talking Sounds made by the audience will be audible to the performer, and provide further feedback. Often when performing juggling, hearing the audience sounds is more important than actually seeing them, as you can hear sounds while concentrating on a hard trick.

4.6.2 How do they see things?

In addition to noting *what* people see, there are several subtleties in the ways in which the performer, audience and computer perceive things.

Computer The Eye-Balls system has a well defined set of perceptions, it can track up to 8 balls, and the hand and head positions of a single performer. These are absolute, in that if it can see 8 balls, it knows the positions of all 8 at once. However, there are limits to these perceptions, described below. Other than these perceptions, it cannot directly detect anything else.

Range: Eye-Balls works on a single camera and cannot track balls that are outside the view. It does however predict the position of balls thrown off the top of the screen, this allows high throws without losing tracking as long as the balls are only affected by gravity once they are off the screen.

Light: Eye-Balls can only track balls and people that are in the view of the camera. If it is dark, glowing balls must be used; in this case the head and hand tracking is turned off.

Occlusion: Because it is a two dimensional system, sustained occlusion can cause Eye-Balls to lose tracking (it will keep tracking a ball using predicted position for a short time). Moving two balls so that they are on top of each other in the view may cause tracking to be lost on one until they are separated, meaning it will think that there is just one ball in that position. Also, if a ball is carefully held in the hand and the other hand used to cover it, it can be moved so that it disappears to the system. This effect is exploited when Eye-Balls is used for 'screen painting', in order to allow a gap to be created in the line being drawn.

Angles: Once the user turns at extreme angles to the camera, the arm tracking may fail, and depending on their hair colour the head tracking may also fail (the head tracking relies on skin colour, blonde hair is typically a relatively good match, but darker or red hair may be less good). Arm tracking typically fails once the arms start occluding each other. The ball tracking is also less reliable when the juggler is turned sideways and the balls are occluding each other.

Catches: The tracker interaction model makes assumptions about the interaction between arms and balls. In particular it assumes that balls are caught by hands. This typically works very reliably, however there is one situation when it breaks, which is when balls are caught on another skin coloured part of the body, such as the elbow. This can cause the arm tracking to briefly think the hands are at the elbows, or the ball tracker to fail to detect that the ball is caught. However when the ball is caught in the hand of that arm again, the arm tracking recovers.

Speed: Eye-Balls tracks at either 25 or 30 frames per second (camera dependent). In some moves, very fast throws are used from hand to hand. In order to measure the possible speeds of inputs, a repeated sequence of these throws was tested across a known distance of approx 30cm. The throw rate in this testing was 3.3 per second, which averaged out to 1.98 m/s, or 7.92cm per frame. This does not take into account the deceleration inherent in the catch and throw back used in this sequence, in practice it is relatively easy to throw between the hands so fast that the same ball in two consecutive camera frames is approx 10cm apart. At this speed, the system has real problems tracking the catches, and sometimes briefly loses the ball. Similarly with throws going high and fast off the screen, a throw to 3 metres will mean that the ball moves up to 30cm per frame which means that the prediction as to where the ball is when it is offscreen may have a large error. This causes the ball to be recognised as a new ball when it returns to the view rather than a returning ball.

Depth: The tracker is only 2 dimensional, so it is only able to make limited estimates of how far forwards or backwards from the camera a ball is by looking at the size of the tracked object.

Performer The performer has more complex perceptions than the computer system. They are typically aware of their body position and the position of the balls (except in the case of drops, when the performer accidentally misses a ball). They will also be aware of sounds and visual cues from the audience. As well as this, they are able to see the monitoring screen showing output from the system, and hear any audio outputs from the system, which allows them to interact with the system. However, there are some limitations placed on their perceptions.

Attention: The performer's attention is divided in three ways, between the juggling they are performing, the monitoring screen (and audio outputs), and the audience. Visually, when the performer is doing a more difficult trick they will have to devote more of their attention to the juggling. This means they will not be able to simultaneously perform hard tricks and watch both the audience and the monitoring screen. When looking at the audience they may find it hard to watch the monitoring screen.

Angle: Depending on their position, the performer will lose sight of the monitoring screen (and potentially of the audience). For example if balls are thrown up and the performer pirouettes beneath them they will be unlikely to be able to see the monitoring screen.

Viewpoint: The performer's view of the performance is very different to that of the audience and computer system which both have a roughly similar view of the juggling. While a good performer will envisage his act from the audience's viewpoint and the monitoring screen may make up slightly for this, they will not be able to see either the audience or the system's view of their performance.

Awareness: The performer will generally have a good idea of what is going on with the ball pattern, however it is clearly not the same as the tracking system, which knows exact positions for each ball at any time. When juggling, a performer does not have the same conscious knowledge of every ball position, they are much more reliant on memory of patterns and visual cues to allow them to see which balls need catching next. They do however typically have a good knowledge of the overall trajectories of the balls in a pattern - evidenced by the fact that many good jugglers can juggle simple 3 ball patterns blindfolded.

Audience The perceptions of the audience are again less exactly defined than those of the computer. They are able to see the performer and the balls, plus the main output screen. They can hear audio output produced by the system and any sounds the performer makes. They aren't able to see the performer's monitoring screen. These perceptions are also limited in various ways:

Attention: In a system with a projection screen and a person juggling, the audience's visual attention is split between the two, they will be unable to concentrate on both at once. This is more of an issue if the audience are very close to the juggler and the screen, in a large scale stage performance the audience may be able to split their attention between both the juggler and the screen at once.

Viewpoint: The audience's viewpoint means that they can only see one side of the performer. They are also often some distance from the performer and screen, limiting the level of detail that they can see.

Awareness: Whilst the audience will be able to see the juggling, they typically are significantly less aware of the details of the performer's actions. For example, even after watching basic juggling of 3 balls, a typical non juggler does not understand the pattern (this fact can be easily ascertained by demonstrating 3 ball juggling to a non-juggler, and then asking them to repeat it; almost universally they try the wrong pattern). Because of this, when juggling to a non-juggling audience, performers often prioritise very obvious tricks, like pirouetting whilst throwing the balls high in the air, or catching with the hands facing downwards, i.e. tricks that make very stylised and obvious patterns of ball or body movement. It is important to remember that this isn't universal, as some audiences or audience members may be more juggling aware, so able to understand some of the more complex tricks.

4.6.3 Foreground and Background Interaction

During development of the juggling tracker, it became clear that a single fixed augmentation of the juggling (such as just displaying coloured trails showing the ball pattern) would not be suitable for most performances. This meant that performers needed to be able to alter the current state of the system during a performance, either to alter how the current display is happening, to change the display completely, or to something in between. In some cases a performer may wish a display to slowly change from one type of augmentation to another, and may even wish to control the rate of this change.

One example of these control movements is in the ball colour script shown in Section 4.3.2. Whilst the performer juggles 3 balls, it does one thing, showing quickly fading trails where the balls are. Then, if the performer puts down 2 of the balls, it causes the system to change state, into a state in which trails fade very slowly, so the performer can write things on the screen.

This control of the system is handled within the juggling tracker itself, using the same tracking as for the creation of outputs of the system. This design aims to make the system as flexible and usable by real jugglers as possible. It is more suitable than the alternative of using external control for several reasons:

- It provides several expressive and potentially continuous modes of input that are not limited in the same way as 'scene change' or 'patch change' buttons in other similar systems.

- It avoids the use of custom or non-standard technology, which is not a possibility if this is to be used by typical juggling groups.

- It does not constrain the control of the system as other hardware interfaces would, for example foot switches restrict stage movement, heel or knee clickers restrict the person's bodily movement in a way not acceptable to jugglers, and hand controlled interfaces are not possible to use whilst juggling.

- It does not require assistants, which are often not a possibility for juggling performers.

This multiple level control system is similar to Buxton's foreground and background interaction [22]. The interactions causing the system to change state are his foreground interactions, where the user is explicitly commanding the system. At other times, the system is simply responding to what the person is doing in a way similar to Buxton's background interactions. What is innovative about the way they are used in the juggling tracker is that both background and foreground interactions are being done in the same tracking system.

Many of the most interesting challenges of performance design for the juggling tracker came from this two level control method. Performers have to work to make the control interactions fit into the performance itself which may require completely different interactions for different performances or styles of performance. The next section describes ways in which the differences in visibility and understanding made a useful contribution towards making control of the system work, and strategies exploited in performances.

4.6.4 Control, Visibility and Attention

The audience's attention may at times be fully focused on the juggling, at other points be watching the screen, and sometimes keeping an eye on both. By taking this factor into account performances can be designed to manipulate the audience attention, using a range of strategies described below. Integrating control interactions into performances requires a good understanding of these strategies as it is often desirable to either draw attention to, or remove the attention from the control of the system.

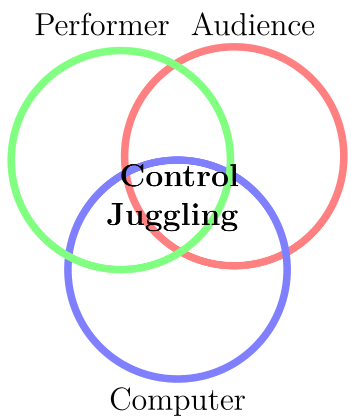
These strategies emerged during development of the tracker, both from feedback from workshops and people testing the system and audience members watching the system, and also from gaining experience in using the system for live performance. Several of the strategies that emerged involve an element of directing the audience's attention, either towards something important, or misdirecting away from things that they are not meant to see. These magic-like strategies make it clear how useful this kind of audience manipulation can be for a performer.

Obvious Control During juggling performances, performers will pick up and put down objects at various points. Often a performer will do some very complicated tricks with three juggling balls (or even two), and then do some simpler tricks with larger numbers of balls (as the larger numbers are harder to juggle). As they change, they attract the audience's attention to the number and type of objects they are about to perform with. These changes in ball numbers are easy to track and can be very effective, as they provide a natural break in the performance. As a natural break in the performance, the fact that they may control the system is very natural, so it does not matter where the attention of the audience is.

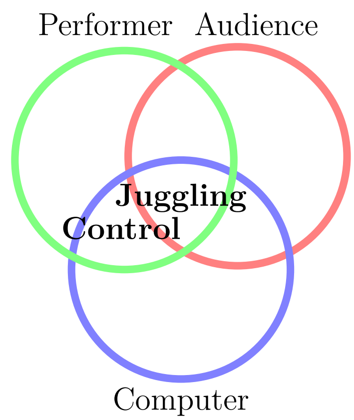
Recovering from Mistakes While increasing the number of balls may be a good control strategy to create obvious changes in sound and music, in some cases, control by detecting a reduced number of balls may not be so useful. This is because of the potential for mistakes in juggling, particularly drops, where a ball is dropped to the floor. In early testing of the system, it was suggested by jugglers that having the system effectively highlight a drop by suddenly changing, was not a good idea.

How jugglers handle drops in performances varies. There are three common strategies. The first is to talk your way out of it by using a 'drop line', a funny line designed to make a joke out of the drop - for example "*I'm just getting back to basics, folks. This was the first juggling trick I learned".* A different but similar strategy is by performing a kick-up to recover, using the feet to kick the ball back into the pattern - this looks impressive, and exploits the audiences lack of knowledge of what the performer actually meant to do. This recovery is very similar to the way magicians use their audience's lack of knowledge of the trick to recover from the failure of part of a trick [99]. Finally, some jugglers will simply tell the audience beforehand that juggling is hard and that they might drop sometimes, and then just pick up drops and continue straight away.

In terms of the system control, different things are required for these two strategies. In the drop line strategy, it is not a problem (and can actually be useful for comedic effect) if the system grinds to a halt on a drop, however there must be a quick way to get it back into the required state. If the magician is going to ignore the drop or kick back up, it may be better for the system to continue audio or video accompaniments. It is even possible to implement a strategy where the screen and audio accompaniments are actually intensified in the event of a drop



(a)Real



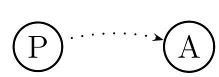
(b)Desired

Figure 4.3: Hiding the Control Interactions in Juggling

in order to direct attention away from the juggler while they restart juggling. This is essentially manipulating the system in order to deceive people into thinking that it was intentional.

Hiding the Control Interactions in Juggling One strategy is to hide the control interactions within the juggling. In terms of perceptions, what is actually happening is all the actions are being done in view of the audience member (Figure 4.3(a)). What this strategy intends is for the system to appear to the audience to just change without them being able to see any specific control interactions - creating in effect the set of perceptions shown in Figure 4.3(b). This is a useful strategy in cases where the audience member is expected to be watching the juggler, but where we do not wish to obviously control the system.

This strategy is typically achieved in the juggling tracker by creating controls based on juggling tricks. This is a direct deception from the performer to the audience member (see diagram on the right). These can be woven into the juggling at times when a control change is required. Often, an audience will not notice subtle differences in tricks - for example, a very simple juggling pattern, the cascade, can be juggled in two different directions (see Figure 4.4). The difference between the two directions is not obvious to many audience members - but is very clear to a computer. This means that a quick change to reverse the cascade can be used to as a control interaction without raising suspicion from



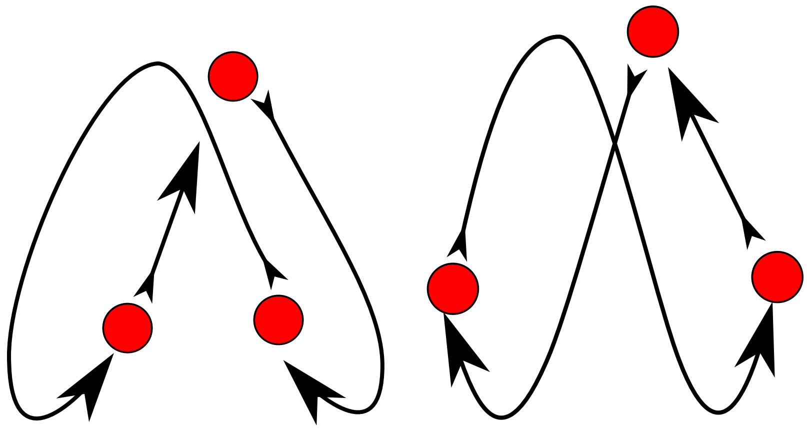


Figure 4.4: Cascade and Reverse Cascade Patterns

Figure 4.5: Moving across stage as a control interaction

the audience.

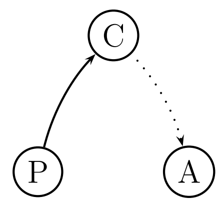
One other simple, non-suspicious control interaction is movement from side to side on stage. By moving across as they juggle the performer can create a continuous input to the system (Figure 4.5). Similar effects can be introduced by moving the juggling pattern up and down.

Expressively Performing the Control Actions A different approach to control interactions is to perform elaborate movements which obviously cause the system to do something. This approach may still use simple underlying actions which cause the control changes, but they are concealed within a larger non-juggling action. As an example, the wizard pose, with both arms holding balls above the performer's head has been used as a control in several performances (Figure 4.6). What the computer is actually tracking is simply the frame shown with the black background in Figure 4.6). However, by making a performance out of the pose, casting a spell, shouting incantations etc., this simple control can be made more impressive, as if the system is responding to something more complex. In one performance with this, the pose was accompanied with the sound of thunder from

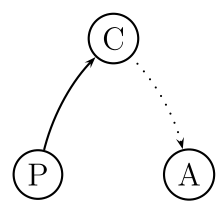
Figure 4.6: The Wizard Pose

the speakers, and the screen flashing white to simulate lightning.

This strategy is essentially using the classic magic trick deception (shown to the right) - using the technology to make it seem that something different is occurring. In contrast to the previous strategy, the deception is being used not to make it seem as if control is not occurring, but rather to make it seem as if control is something that is skillful and interesting.



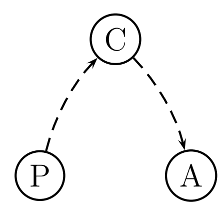
Directing the Audience Attention Toward the System Rather than hide the interactions by juggling, another strategy, useful when a set of control interactions are required, is to create a situation where the audience is not watching the juggler. One way to do this is by using control interactions that make a loud noise and flash the screen. This draws the audience's view towards the screen, and a sequence of control interactions can be performed whilst the audience watch the screen. As well as control interactions, this can be useful for picking up or putting down juggling equipment, or for moving to a particular place on the stage without an obvious break. This type of interaction is essentially using the computer system in order to trick the audience into not watching what the performer is doing (another magic trick deception - see figure on right).



Directing Attention Away from the System In some situations it may be useful to bring the attention back from the screen, and concentrate it on the juggler. For example in the example performance in Section 4.5, for the final big trick the performer wishes all eyes to be on him rather than the screen. Part of this is simply communicating with the audience as in any performance, but design of the way in which the system augments the trick is also useful here. A simple method of directing attention away from the system, is to make the screen show nothing at all, or a static image, such as the text saying *'hardest trick in the world'* in the example performance.

At some points in performance the performer may want the augmentation shown on the display to change abruptly, yet without a jarring break in the performance. In this case, a related approach can be used, rather than the system becoming boring in order to lose attention, the performer becomes more interesting, by performing impressive or fast tricks, moving around stage fast, shouting, or otherwise catching the audience's attention. In this way, the screen can abruptly change, whilst the audience are tricked into not being aware of it. This uses a very similar technique to the misdirection techniques above.

Multiple Jugglers The juggling tracker detects the position of the arms and head of a single juggler. One interesting thing tried by jugglers using the system, was putting multiple jugglers in front of the system. The loose coupling of the hand/head tracking and the ball tracking means that the ball tracking still worked in this situation, although with a loss in accuracy. The presence of multiple jugglers is potentially detectable in the tracking system, however the design of the system was intended to avoid restricting possible inputs as much as possible, so it simply ignores this. This designed-in suspension of disbelief (which can be described by the diagram on the right) allows for interesting comedy tricks to be performed by appearing to fool the tracking system with two or more jugglers performing. A nice side effect of this is that when children were allowed to play with the juggling tracker at an early public demonstration, two or three of them could have a go at one time, which allowed children who could not juggle to have a fun experience with the system.



4.7 Conclusions

In this chapter, the Eye-Balls juggling tracker and the design process used to develop it has been described. The development of this system was a major inspiration for the ideas of the use of deceptive techniques, directing attention, and misdirection. In particular the challenge of controlling the system whilst concurrently juggling requires particular thought. This has demonstrated how techniques used to achieve a successful and smooth performance with the system rely to a large extent on directing the attention of the audience between the outputs being created by the system and the juggling itself, both of which are important parts of a good combined performance. An important point demonstrated here is that within a combined or augmented performance, there may be times when one part of the performance is more important than the other, or when one part of the performance is doing something that is outside the frame of the performance. In a typical setup, it is not possible to physically remove either part from view at this point. Instead, attention may be drawn away from one aspect of what is occurring, for example the performer may wish to have the audience's full attention at times, and at others, he may wish for the audience to be looking at the screen and not to see what he is doing. This is in many ways similar to the misdirection and attention direction techniques used by magicians.

This chapter has demonstrated the usefulness of this kind of magic inspired technique in an augmented performance situation. However, given the real world performance nature of this system, more nuanced exploration of these deceptions within the system is very difficult. In the next chapter a system designed specifically using these ideas is presented. This system allows exploration of the processes of misdirection and attention direction within a smaller scale performance that is easier to analyse in depth.

Chapter 5

The Cup Game

The Cup Game is a computer augmented version of a classic magic trick and con-game also known as The Shell Game or Find the Pea. The traditional version of this trick uses three cups (or walnut shells, sea shells or bottle tops), and a small object, traditionally a pea or bead. The object is placed under one of the cups. The cups are shuffled, and an audience member is asked to guess which cup the bead is under. Invariably despite careful observation, the object will not be under the expected cup. In the case of the con-game version of the trick, people are tricked into losing significant amounts of money by groups of skilled con-artists.

This computer augmented version of the cup game is designed for several related tricks in addition to the conventional version where the performer shuffles the cups to trick the audience. It also allows the audience member to try and trick a blindfolded performer and also has a beat-the-computer version where a person has to try and beat a computer designed to watch the movement of the cups and detect when it is being tricked.

The technology used in the augmented cups game is based on a similar computer vision algorithm to that used in the juggling tracker, with a custom built table mounted cabinet to mount the camera and show the trick.

This chapter documents the successful development and performance of a routine using this system. This explicitly deceptive setting gives further insight into the use of deception in performance which is backed up by quantitative and qualitative analysis of the application of the performance to a number of participants. The development is carried out in an iterative process, informed by the deception framework described in Chapter 3.

5.1 Why?

The cup game is included here as it explores several distinct classes of deception in a very controlled performance setting. In particular, it focuses on the use of a system that is explicitly designed for deception, both to deceive others, and to be deceived itself (or to catch deception). It is designed for close-up use, with one performer and one audience member. This makes it significantly easier and quicker to iterate and evaluate performances than is the case with a one to many performance system. Due to the physical setup in which the performance occurs, it is easy to record video of audience, performer and computer views of the performance, in order to analyse them after the performance is done.

In this chapter, the development of a performance of the cup game is documented. The initial iterative development of this was done with colleagues from within the School of Computer Science, University of Nottingham. This performance was iterated with new participants until it reached a steady state, where no alterations had been made to the performance for 5 participants. Once this steady state was reached with these internal participants, the performance was evaluated with outside participants.

The character of the performance shown here is such that each deception strategy is separate, and it is clear whether it has succeeded or not. The analysis of the performance allows the capture of both numerical data, demonstrating as to the success of the deceptions within the performance, and also qualitative data relating to ways in which people behave and feel during the performance. This allows us to explore how the deceptions work in a performance, and the possible effect that various attributes such as technical knowledge have on this.

5.2 Related Performance Work

Throughout the history of magic, magicians and others using similar techniques have integrated modern technology into their performance. Even in ancient history there is evidence that Egyptian priests used mirrors and speaking tubes to produce mysterious effects that were attributed to gods [129]. Robert-Houdin famously used electro magnets (at the time cutting edge technology) to make a chest seem very heavy and then suddenly become very light [111]. Computers are an extension to this tendency to use the most modern technology .

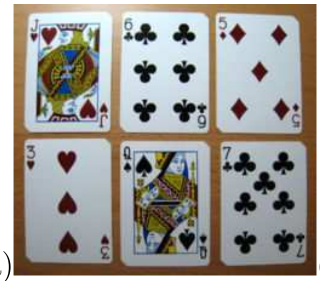
(a)  (b) 

Figure 5.1: The Missing Card Trick

One of the most famous tricks involving technology is Penn and Teller's "World's Most Expensive Card Trick". In this trick Teller was shown in the studio controlling 2 computers; one which supposedly analysed a video of a pack of cards fanned in front of a camera, and another which controlled a large public screen - in Picadilly Circus in London, and Times Square in the New York version of the trick. Penn was shown in the street where the screen was, choosing a person off the street to buy a pack of cards and pick a card out in secret. The rest of the cards were shown to the camera, and then the missing card was displayed on the big screen, behind the head of the person but visible to the TV show audience. What in fact happened was that the person choosing the card was an accomplice, and knew to always pick a particular card, so all the technology was just a front [103]. This is similar to Robert-Houdin's claim to use anaesthesia in this shows [129] - where he used a new and seemingly magical technology as a blind for what is actually a standard trick.

Several online card tricks have been created, most of which are versions of the classic missing card trick, in which a set of cards are shown and the audience member told to choose one (Figure 5.1(a)). Then the cards are shown again, and the audience member discovers that their card has mysteriously disappeared (Figure 5.1(b)), and that the performer, or in the online case the computer appears to have read their mind.

As well as this common trick, performer-controlled online tricks have been released that use a secret code between the computer and the performer to run the trick. For example, the iPolygraph trick by Jim Bumgardner [142] shows a 'continue' button, which invisibly is split into multiple sections, allowing the performer to pass the number and suit of a chosen card to the computer for it to pretend to guess (Figure 5.2). This trick uses a similar secret communication

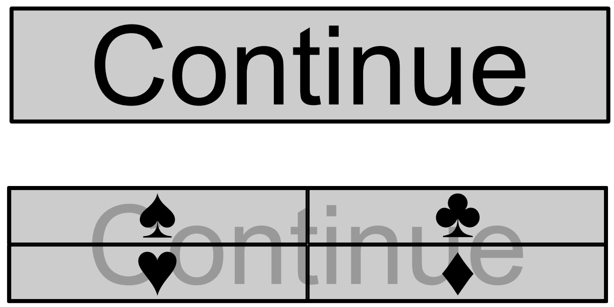


Figure 5.2: Secret button code used in the Psychic Computer trick

process to the Blank Card Trick described in Chapter 3. There are various other versions of these ‘click in the right place’ tricks available online.

Assistant tricks, such as Jason Litchfield’s Mother of All Online Card Tricks [149] (described in Chapter 3), use an assistant working over a network or phone connection to tell the computer which card to pick.

Tenyo, the Japanese magic company that developed the Magic Made Fun computer game (see Chapter 3), have also made many tricks designed to run on mobile phones including versions of all three of the above tricks, and other tricks using technologies such as speech recognition. These are all in Japanese, and no English version of the tricks is available.

Magic as a metaphor has been used in several HCI projects such as the Magic Table [11], a table which performed tasks such as automatically digitising things written on sheets of paper left on it, and the Magic Book, a physical book which is augmented using a see through pair of display glasses to create a 3d view on top of the physical book [12]. These are not claiming to really be magic, rather are describing the technology as being similar to magic, creating an effect similar to having a magic table that remembers things written on it or a magic book that is a portal into real 3d worlds.

An interesting side effect of the way that the Cup Game is performed is that by taking part in a performance an audience member learns about the computer vision technology employed and the strengths and weaknesses of the system. In the past, performers and educators have used magic as a tool to teach other aspects of computer science and mathematics [32].

5.3 Technology

The technology developed for the Cup Game is a combination of a physical cup table which houses the camera looking down onto the cups, and a laptop, which is below the cup table running software which visually tracks the cups. The physical setup is designed so that the technology can be either hidden or revealed from the audience or performer with the use of sheets of blackcloth.

The physical setup is designed both as a practical way of mounting and hiding cameras and computers and also to frame the performance aesthetically. It is designed to look reminiscent of a miniature theatre and is run in a small dark room, with the only light coming from within the cup table itself, giving a dramatic feeling to the setting and focusing the audience on the cups themselves (Figure 5.3).

Figure 5.3: The Cup Table

5.3.1 The Cup Table

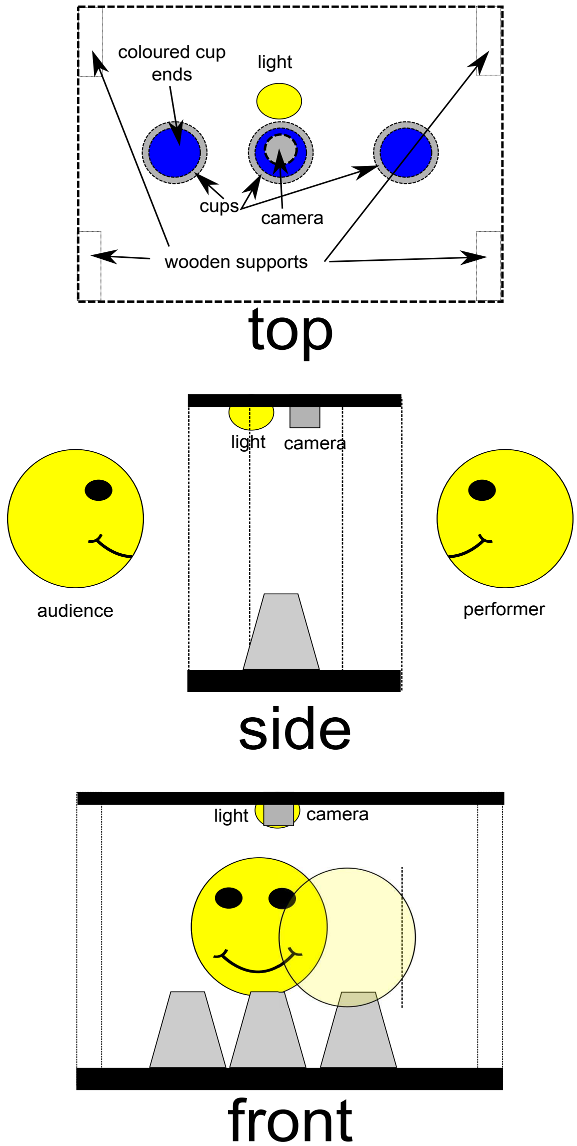
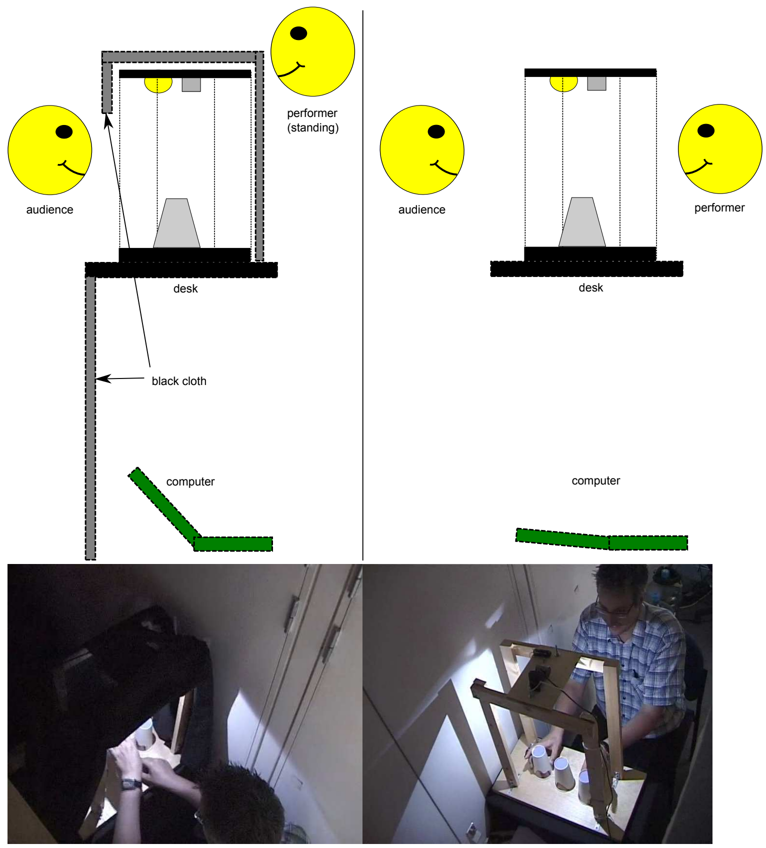


Figure 5.4: Schematic of the Cup Table

The cup table is made of two sheets of wood, a thick base which serves as a table for the cups to be moved on, and a thinner top sheet which is used to mount a web camera directly above the surface the cups are moved on. These two surfaces are connected by a simple wooden frame. There is also a small high-powered LED embedded in the top sheet, which is used to light up the cups.

This construction is placed upon a desk, with a laptop put on the floor below, which is running the video tracking software.



(a)Hidden (b)Revealed

Figure 5.5: Hiding and revealing the technology

Hiding and Revealing Technology In one of the tricks in the performance the existence of the technology must be concealed from the audience member. At the same time, it must be made obvious to them that the performer cannot directly see the cups. 2 blackcloths are used which cover firstly the top and back of the cup table, and secondly cover the front of the desk, so the audience member cannot see the laptop (see Figure 5.5(a)). The camera lens is made harder to see by a small flap of blackcloth over the top front edge of the setup. It is also right next to the bright light which makes it hard for the audience member to look directly at it.

When the technology is revealed, the blackcloths are removed, and the laptop screen folded down. This setup is designed so that both audience and performer can see one screen and so the tracking images on the screen are correctly aligned with the real movement of the cups (Figure 5.5(b)).

The space the audience is sat in is deliberately narrow in order to avoid them

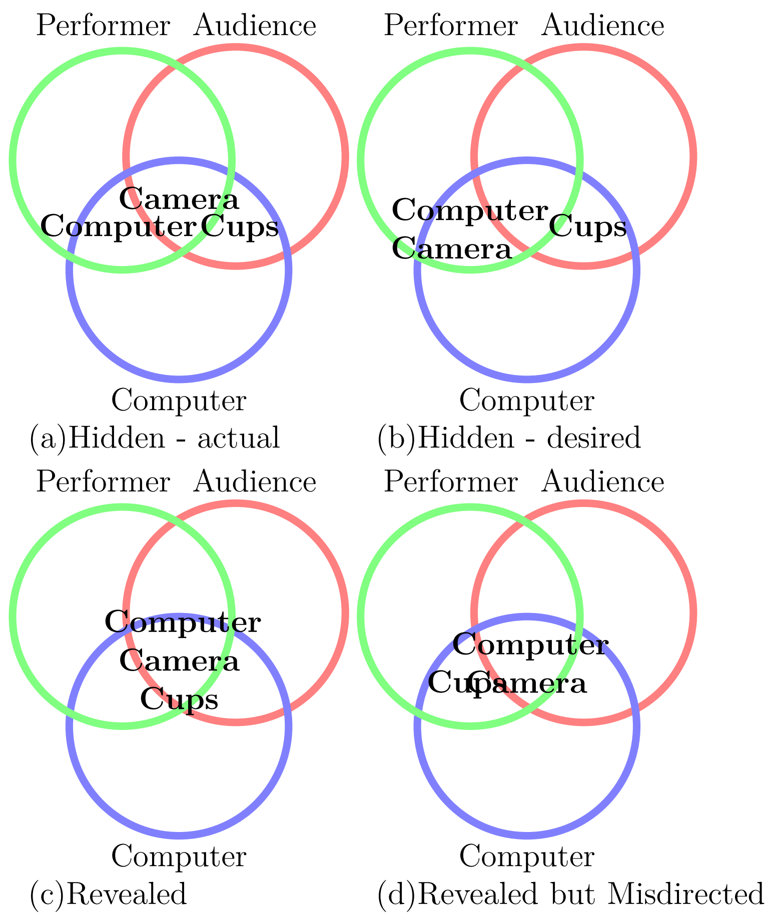


Figure 5.6: Perceptions and Hiding / Revealing the Technology

looking around the side of the table or wandering round to the performer's seat before the performance starts.

In terms of perceptions, in the hidden state all that is actually visible to the audience member is the cups and the camera. The performer can only see the computer. The mapping between cup movement and camera tracking / computer display is only understood by the performer (the audience member does not even know the computer exists). The actual physical visibility is shown in Figure 5.6(a). The design directs the attention away from the camera, so people don't notice it, creating effectively the situation in Figure 5.6(b). When the table is used in the revealed state, everything is visible to everyone (Figure 5.6(c)). However, in one of the tricks misdirection is used to direct attention away from the physical cups, creating in effect the situation in Figure 5.6(d).

5.3.2 Software

The cup tracking software uses a simplified version of the ball tracking algorithm used in the juggling tracker. The tracker can be simpler because of the way the cups and the cup table are designed. The cups are tapered towards the bottom, with the flat bottom being coloured so that the tracker can see it. Due to the taper, the coloured bottoms of the cups can never touch - this means that the tracker looking from above can always see separate blobs of colour for each cup. In normal use the cups can never occlude each other and due to the shape of the cup table, cannot be moved out of the camera's view. There is also a fixed number of cups which are assumed to be on the table at all times. The fixed number makes tracking of cups occluded by the hands easier - they can always be assumed to be still somewhere on the table; the cup game software assumes unseen cups are in the same place they were last seen until a cup reappears. This is in contrast to the juggling tracker which assumes balls have been dropped or taken out of view once they have not been seen for a short time and has algorithms to detect the juggler changing the number of balls they are using. In one state in the performance it has a 'cheating alarm' designed to detect trickery. This is set off when the cup tracker cannot see a cup for a very long time and is based on the assumption that it has been covered, removed from the table, or the light has been turned off.

The tracking is of the cups and not of the bead; this means that the tracker will only work if the bead actually remains under the same cup. This is key to the functioning of several of the tricks.

The visualisation of the cup tracker is also simplified compared to the juggling tracker. It simply displays a combination of the video image from the camera, coloured dots to show the tracked positions of the cups, and text instructions. A simple state machine based controller is used in order to run the visualisations in the order required for the performance.

5.4 The Performance

This section presents a performance developed with the cup game setup. This is a one on one performance involving a set of short tricks. Each one is designed to explore the creation of a certain effect, by performing a deception on the participant, deceiving the computer system, or allowing the participant to try and deceive the computer system. The tricks are designed to be run as a single unit which tests a variety of ways in which tracking technology may be used in performance. The performance is also aimed at demonstrating to the participant both the existence of computer tracking of objects, and also some non-obvious aspects of the tracking technology, particularly how it is limited by the assumptions that were made when it was created (most importantly the fact that it assumes the bead always stays under the same cup). Thus it both entertains and educates the participant. Analysis of the performance being run provides useful data as to the potential success of these deception techniques.

The performance was developed through an incremental process. It was initially piloted with members of the Mixed Reality Lab, University of Nottingham. These were all computer scientists, and as such, had some knowledge of the computer vision technology involved. Using these technologically knowledgeable users as initial test subjects was useful, however due to the large percentage of programmers (80%) and people who have used computer vision technology (90%) in the study this type of user did not provide suitable insight into how the average person would interact with the system, so a further test was run with external volunteers, recruited using a poster placed at bars, cafes and places around the university where students socialise (Figure 5.7).

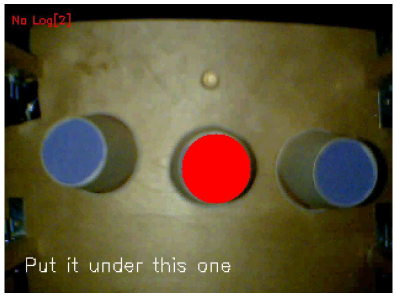
The visualisations for each trick use a selection of the 5 screens shown in Figure 5.8. They are always shown in the order described, although some stages are skipped in some tricks. (The trick descriptions refer back to this figure.)

The performance is split into 6 stages - a brief introduction, three tricks which are successively revealed, a big reveal demonstrating the weaknesses of the tracking technology and the ways in which these have been exploited in the previous tricks, and finally a short set of questions about the background of the audience member. The stages are described here in the order they occur in the performance.

In each section, what occurs in that stage of the trick is described. This is followed by a description of the effect that the performer is aiming to achieve,

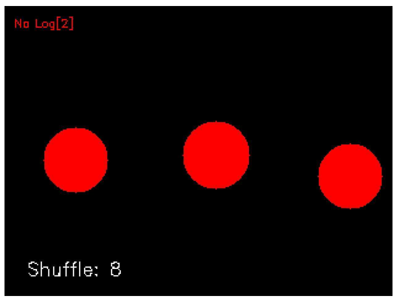
Figure 5.7: Magic Experiments Poster

1



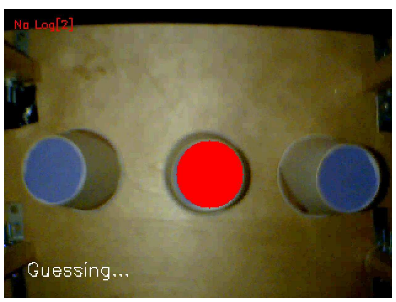
The visualisation instructs the user to put the bead under the central cup, with video of the cups shown, and the middle cup highlighted.

2



The video fades out, and the user is instructed to shuffle the cups - with red dots highlighting the tracked position of each cup, and a countdown for 10 seconds, after which the next stage is run.

3

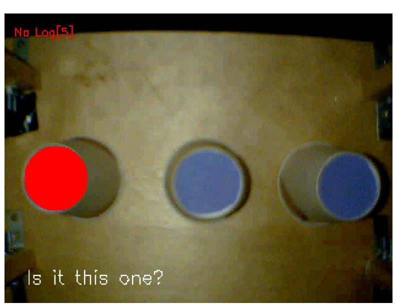


The text says 'Guessing' and the red dot flickers over each of the cups randomly for 2 seconds

4

The text says 'I think I know' (which one it is).

5



'Is it this one?' - The computer highlights its guess.

Figure 5.8: Visualisation of the Cup Game

and the deception strategies which are used. This is followed by a set of research questions which are designed to inform the analysis of that section. These are split into two types, questions where the answer is a simple categorical piece of data, gained either from watching the actions of the audience member, or directly from a question asked after each trick, and more open ended qualitative questions. The questions do not exactly correspond to actual questions asked of the audience member - for example *did the audience member pick the right cup?* type questions can be answered during the performance. A full script of each section is in Appendix B.

The tests were recorded using two video cameras showing over the shoulder viewpoints of the audience and performer (and the face of the person on the other side), plus video recorded from the web camera on the table and a recording of the tracking state of the software. This gave a rich set of video and audio data, which could all be played back in synchronisation to see the state of the system at any point during a run. Answers to the questions were transcribed at a later date.

All questions were asked directly by the performer and written questionnaires were not used. This is because the questions needed to be asked at several points during the test, as the revealing of previous stages in the tricks would have an effect on how people answer the questions about these stages. From a performance viewpoint, stopping the performance after each stage to write down the answers to questions would not be sensible.

5.4.1 Stage 1: Introduction

The audience member comes into the room and sits down in front of the cup table, which has blackcloths covering the side towards the performer and hiding the computer from the audience member. At this point, the performer is standing. The audience member is introduced to the cup game, and shown the cups and the beads, and told how the trick conventionally is run, with the performer shuffling the cups and the audience member guessing which one has the bead underneath it.

5.4.2 Stage 2: The Cups by Sound

In this trick, the performer explains that he is going to let the audience member shuffle the cups, and then by listening to the sound of the cups being moved, detect where the bead is. The audience member is given two instructions, firstly that when they are moving the cups, to hold them by the sides, supposedly to improve the resonance of the cup making the sound, and secondly to only move one cup at a time, to make it possible to hear the movement.

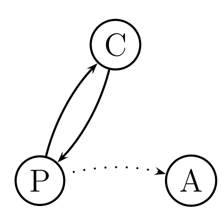
The audience member cannot see the computer or the camera at this time, so does not know that there is any technology involved. When sitting down, the performer can see the screen, which is showing visualisation 5, just tracking the cup which is initially in the middle.

However, to make the trick less easy to guess, the performer does not watch the hidden screen, instead they stand up and put on a blindfold, and face away from the screen whilst the audience member shuffles the cups. Once the audience member has shuffled the cups, the performer announces that he knows which one it is under, turns round and sits down, taking off the blindfold. A quick glance at the screen now allows them to get the correct answer.

After the trick has finished, the audience member is asked a short set of questions as to what they thought happened, and how they thought the performer did this trick. These allow the questions described below to be answered.

Desired Effect The desired effect here is that the audience believes the performer can tell which cup the ball is under in some magical or highly skillful way, as described by the cover story.

Deception Strategies Used This is simply using a computer as a tool, to help in a deception played directly on the audience by the performer. The audience member doesn't know that a computer or camera are being used yet. The computer is hidden under the table, and the camera is not visible due to the blackcloth.



The deception is reinforced by the invention of a cover story, which is designed to make them think about the sound of the cups rather than any visual tracking. The instructions given in the cover story are also designed to make the person move the cups in a manner which is easily tracked by the computer.

When giving the final answer, the performer says 'I know which one it is under', whilst still blindfolded - yet only says the answer once they have removed the blindfold and seen the screen. This is again designed to make the person think that the discovery happened earlier, whilst the blindfold is on, which makes them think that a visual method of doing the trick is not possible.

Categorical Data

- Did the tracking work?
- Do they believe the explanation?
- Did they spot the camera / computer?
- Did they guess there was a computer involved?

Qualitative Data

- If they did not believe the original explanation, what explanation did they come up with?

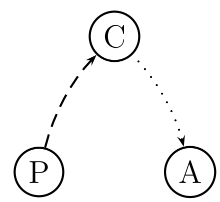
5.4.3 Stage 3: In Front of Your Eyes

To begin this section, the performer reveals the computer and the camera, by removing the blackcloths and by folding the laptop screen down so that it can easily be seen by both performer and audience. Then, purporting to show how the previous trick worked, he starts the visualisations. This time, all visualisations are used. Firstly, the screen tells him to put the bead under the central cup, and he does this. Then, the screen says to shuffle, and he says - '*if you look at the screen now, as I move the cup, the circle on the screen moves. Watch carefully and pay attention to this one* [shakes the centre cup] *which is the one with the bead under it.*' The cups are then shuffled for 10 seconds, after which the computer says '*Guessing*', followed by '*I know which one it is under*'. At this point, the audience member is asked which one they think it is under. If the trick has worked, they think it is under the right hand cup from the performer's point of view. Then, a key on the computer is pressed, to ask it which one it thinks the bead is under, and it shows its guess which is the left hand cup.

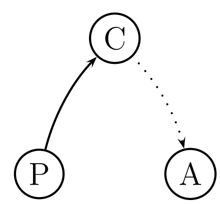
As the two cups are lifted, the computer is shown to be right, and the person wrong. Again, the audience member is asked what they thought happened.

Desired Effect The desired effect here is that the audience believes the computer can tell which cup the ball is under in some way that is better than their ability to watch it.

Deception Strategies Used This trick is actually composed of two deceptions. The first - labelled (a) is that the design of the visualisations deliberately does not show the camera video in the shuffling stage - it only shows circles showing where each cup is. Demonstrating the tracking is used here as a misdirection for the audience, to make them look at the screen, and not at what the performer's hands are doing - which is to lift the cup off the bead, and lift another cup over it. Performing another trick whilst purporting to be demonstrating a previous trick is a classic example of misdirecting the audience as to where a trick begins [99]. This means that whilst the audience have accurately used the computer to track the correct cup, the bead has been moved under another cup, right in front of their eyes.



(a)



(b)

Deception (b) is that when the tracking is started for this trick, a particular key is pressed which tells it that it is being fooled. This causes it to always pick the left hand cup as the final guess - a simple hard-coded guess. The performer, who knows which cup the bead is under, can ensure that the cups end in the correct configuration.

Categorical Data

- Did the screen display work as a misdirection?
- Did the computer pick the right cup?
- Did the audience member pick the wrong cup?
- Do they think that there was a trick, or just that the computer tracked the objects better than them.

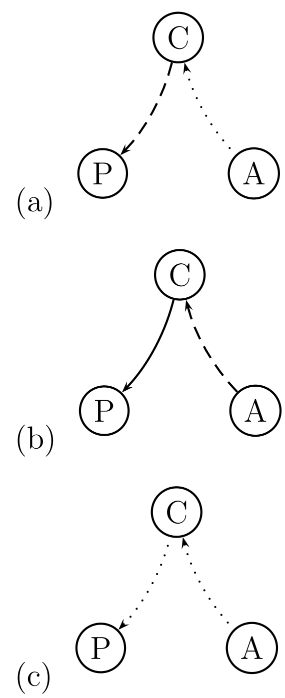
Qualitative Data

- If they thought there was a trick, what explanation did they come up with for it?

5.4.4 Stage 4: Get Your Own Back

In this section, the audience member is allowed to try and get their own back on the computer, by having a go at fooling it themselves. Screens 1-3 and 5 are used in this, with no pause between the shuffling, and the computer guessing which cup the bead is under. This is so that the audience member doesn't think the computer is being helped at all - in case they have guessed how the previous trick was done.

If they have trouble the first time, the performer emphasises that they can do *'absolutely anything'* to trick the computer, as long as it doesn't notice that it has been tricked. This avoids the situation where a person repeatedly tries to do the trick in the same way and fails. In some situations, the computer will notice it is being tricked, and set off an alarm sound, which shows it has detected someone trying to fool it. The person is allowed to try fooling it multiple times, until they manage to trick the computer successfully. This is again followed by questions relating to the strategy they are applying.



Desired Effect The desired effect here is that the audience manages to trick the computer somehow, but that it is not too easy for them to do so.

Deception Strategies Used This is essentially about the audience deceiving the computer system, and causing it to report false results (a). If they fail, or set off the alarm, the trick is following strategy (b) above. However, some audience members treat it as fooling the performer, and may even perform their trick well enough to fool the performer leading to strategy (c).

Categorical Data

- Did they manage to fool the computer?
- How many tries did it take?
- Did they exploit the weaknesses of the computer when trying to fool it, or do they fool it in the same way as they would a person?
- Does having technical knowledge make it easier to fool?

Qualitative Data

- What strategies did they use to fool the computer?

5.4.5 Stage 5: The Big Reveal

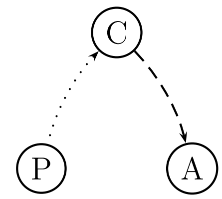
In this stage, the performer demonstrates the trick shown in stage 3 and explains why it works, primarily because the computer is only displaying the cups, and not the bead itself. They also demonstrate the way in which the computer doesn't really guess at all in that trick. This demonstrates the fundamental weakness of the tracker, that it is tracking the cups, not the bead itself.

As well as demonstrating the trick, in this stage, other ways to fool the computer are discussed, and which ones it can and cannot detect. These are:-

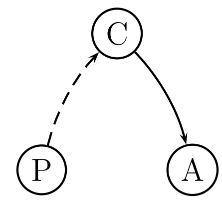
1. Putting the bead under the wrong cup at the start - not detectable.
2. Covering the cups from the top so the camera can't see them and switching - detectable.
3. Turning off the light, so it can't see at all - detectable..
4. Moving the cups out of view - detectable.

Desired Effect Hopefully after this stage, the audience member will have a much better grasp of how the tricks worked, and what the limitations of the computer tracking system are that were being exploited.

Deception Strategies Used In this section, the performer is either successfully fooling the computer, forcing it to display an incorrect answer to the (knowing) audience member (a), or is being caught out by the computer, at which point it communicates this to the audience member (b).



(a)



(b)

5.4.6 Stage 6: Background Information

In this stage, the performer asks a few background questions about the audience member which are designed to get the information described below. They are also asked about their feelings about each of the tricks, and particularly about the revelation of the technology.

Categorical Data

- How often do they use computers? (categorised into daily, every week, rarely, not at all)
- Have they ever programmed a computer?
- Have they previously seen or used computer vision technology? (not seen, seen, used)
- Have they seen magic performed (on TV, or live?)
- Have they ever performed a magic trick themselves? (no, as a child, as an adult)

Qualitative Data

- Did they enjoy the tricks?
- How did they feel about being tricked by the computer?
- Was it a disappointment to discover that there was a computer?
- Was it a disappointment to discover how simple the second trick was?
- How did they feel about fooling the computer themselves?

5.5 Development of the Performance

The performance was altered several times during the initial piloting with lab members, both in terms of altering the script of the performance, and also the visualisations. This section describes the problems encountered during this development process, and how the performance was altered to take account of them.

5.5.1 Tracking the Tracker

After reviewing the first video, it became clear that just video of a trick was not enough to review exactly what was happening in the tracking system. It was not obvious what the system state was at each point, particularly in the case of a trick going wrong.

Because of this, logging of the computer's view and internal state (which visualisation it was showing, where it thinks the cups are), was added to the system, allowing this to be played back in tandem with the video of the performer and audience members' viewpoints of the trick.

5.5.2 People Trust Computers

In the initial version of the In Front of Your Eyes trick (Stage 3), the computer showed an answer and the person was then allowed to choose a cup they thought was correct. However, in practice, people who were even slightly uncertain about which cup was correct would trust the computer's choice over their own. Because of this, a pause was added where the audience member is asked for their answer before the computer shows its answer.

5.5.3 Audience Failures

These are not really audience failures, rather failures of the performer to direct the audience effectively. In Stage 3, audience members in two of the early pilots lost track of the cup completely. This caused problems as the computer's guess of the correct cup gave away the trick. This was also fixed by the pause added above, where the performer checks that the audience think the ball is under a particular cup or restarts if they are not sure. There was also significantly more verbal encouragement to them to keep track.

5.5.4 Following the Rules

The rules set up in Stage 2 in order to make the tracker work reliably for that trick gave people problems when they were tricking the computer in Stage 4 as some of them tried to follow these rules in the later stage. This was fixed with alterations to the script - firstly, one of the pilot participants said that this was like '*beating the computer, getting my own back on technology*'. This idea of beating the computer was brought into the script for Stage 4. As well as this, a cue was given after the first failure to beat the computer, where the audience member was told to do anything they liked to beat the computer.

5.5.5 Software Errors

A bug in the initial version of the software meant that in two of the first 5 tests, the system hung whilst Stage 2 was running. This was a serious problem as having to reset the software gave away the game and revealed the existence of the technology. This bug highlighted the importance of stability for software when the existence of the technology itself is hidden from the audience.

5.5.6 Performer Errors

Stage 2 requires the performer to put the cup containing the bead in the correct place on the table. Achieving this whilst also concentrating on misdirecting the audience member towards the screen was a failure in two of the tests, meaning that both the computer and the audience member chose wrong cups. Whilst the obvious fix to this is to perform the trick correctly, a get out was prepared for the case where this happened, implying that the performer had managed to fool both the computer and the person - this is equally impressive as a trick, since the person does not know what the trick consists of in advance.

5.5.7 Misdirection Failures

In early runs of stage 2 the performer looked down at the screen too obviously. Also, in the first couple of performances the blindfold was not used. In both these situations the audience member guessed the performer was looking at something. The blindfold and the verbal misdirection of saying that they knew which cup it was under whilst still blindfolded was used to make this misdirection more effective. In stage 3 care had to be taken not to look too obviously at the cups, and also to give the audience member verbal cues which suggested watching the screen not the cups, or else the audience member would spot the bead being moved.

5.6 Performance Trial

5.6.1 Demographics

17 people responded to the advert and went through the trial performance, 9 male and 8 female. 14 were undergraduate students, 2 were post graduate students, and 1 was a lecturer.

All used computers regularly. 10 (59%) had computer programming experience. Only 2 (12%) had used computer vision technology, and a further 2 had heard of the technology before, meaning that for 76% of the audience members this was the first they had heard of or used computer vision interfaces.

As for their experience of magic, all 17 had seen magic performed on television and 9 (53%) had been to see magic performed live. As well as this, 11 (65%) had performed a magic trick themselves, mostly as children, although 1 person still performed magic tricks as a hobby and one other had performed stage illusions at a semi-professional level.

5.6.2 Did it Work?

The basic numerical results demonstrate that this design based on the deception framework in Chapter 3 was able to create a successful performance. All participants reported that they enjoyed the performance, with none choosing to stop before the end.

The Cups by Sound and In Front of Your Eyes tricks proved reliable both in terms of the technology working, and the audience members not spotting the trick. Cups by Sound failed once which after analysis of system logs and video showed to be because the user was not following the instructions and was covering the top of the cups. In one run of In Front of Your Eyes, the audience member spotted the switch, so they knew at the end which cup the bead was under and the trick failed. This failure, and the reasons for it occurring are discussed in Section 5.6.3.

The physical hiding of the technology in Cups by Sound also worked well. It was only spotted by one person (who was one of the people who had already used computer vision software).

Beating the computer in the Get Your Own Back stage was a challenge for several participants, with 5 participants unable to beat the computer even given 3 tries at it.

5.6.3 How did it work?

The questions asked of the audience members between tricks and after the performance allow for analysis of what factors made each trick successful, and how these combined in the performance. These are split into several themes, which are addressed in the following sections.

Framing Hidden Technology The framing of the Cups by Sound trick, with no technology visible and the only visible parts of the trick being made of wood and cloth, clearly led people to think that there was no complicated technology involved. In a slightly unexpected result, 10 participants believed the cover story of listening to the sound. Of the 7 who didn't believe the cover story, only one guessed that it was something electrical as they spotted the camera in the top of the cup table. The other non-believers guessed other low technology things, several of which were clearly inspired by the common perception of magicians, "*maybe there is a mirror somewhere*", "*something in the table*", "*I have no idea...*". It is clear that rather than actively think that the performer was doing the trick in a particular way, participants simply thought that the cover story was unlikely - as described by one participant *"it is feasible that somebody might have a really good sense of hearing but quite improbable."*

The use of low-technology materials for the visible parts of the trick and framing with a very different but believable cover story led to one of two outcomes, either that people completely believed the cover story, or secondly that they didn't believe it, but did not know what was really happening. The trick was successful as it was either pulling off a rare feat of skill, or performing an inexplicable trick, both equally good outcomes in this case. The performer must be prepared for either outcome (magicians would argue it is also useful to have a backup plan ready in case the audience member completely sees through the trick [99]).

Tries to Beat Computer	**1**	**2**	**3**	**Gave Up**
Number of People	2	3	7	5

Table 5.1: Number of tries to beat the computer

Times Cheating Alarm Set Off	**0**	**1**	**2**
Number of People	6	10	1

Table 5.2: Cheating Alarm Incidence

People Trust Computers During the In Front of Your Eyes trick, people were asked whether they were sure that they'd picked the correct cup, most of them were certain at this stage that they had the right cup. However, after the computer had shown them the correct cup, 10 people (59%) believed that they must just not have been as good at following the cup movements as the computer. Of the others 2 thought that the computer screen must have been showing a different set of movements to those happening above, and 4 thought that there must be a trick, but were unsure as to what it was (1 person spotted the trick). This trick revealed a high level of trust in the computer, to a level that roughly 60% of people were willing to change their mind about their ability to visually track objects when the computer beat them. This trust of computers is demonstrated by the exclamation of a participant in the final reveal section - *"I didn't know computers could cheat."*

Tricking the Computer is Hard Table 5.1 shows the number of tries that each person took to beat the computer system. Only 2 people (12%) managed to beat the computer on the first try, and 5 (29%) actually gave up after 3 tries. The majority of people took 2 or 3 tries to beat the computer. For all except for the 2 who beat the computer first time users were prompted that they could do anything they liked to beat the computer. Each strategy the user tried was categorised as being either a strategy they would use to beat a person, or something specificly adapted because they knew they were trying to beat a computer. Arguably, this is affected by the framing of the situation - as can be seen in Table 5.3, the prompt,

Computer Specific Strategy	**Before Prompt**	**After Prompt**
Yes	3	9
No	14	7

Table 5.3: Use of Computer Specific Tricking Strategies

given after the first failure, altered the strategies that people used to try and beat the computer, with only 3 people trying specifically to beat the computer to start with, and 9 (out of the 15 who took more than 1 try) adapting their strategy to the computer after the prompt. This effect is significant with P=0.031 (Sign Test). The majority of people in their first try just tried to move the cups very quickly (13 people, 76%), following the method that had been used (by them and also by the performer) in the preceding tricks.

Those adopting a computer specific strategy typically first tried to cover the top of the cups, setting off the cheating alarm, and then tried other strategies. The cheating alarm was set off at least once in the majority of cases - mostly by people who covered the tops of the cups for too long. 2 people also tried stacking cups on top of each other, which again set off the alarm. Those who did not adapt their strategy to the computer typically just tried to move the cups quickly, as if they were trying to trick a person.

Only 3 people thought of the strategy of changing the cup which the bead was under rather than simply trying to move the cups quickly. Most of the others succeeded by accidentally very briefly covering the tops of the cups (too quickly for the tracking system to detect it was being fooled), or by doing other things which caused a failure due to tracker errors, such as shuffling the non-bead cups to distract the computer, moving cups together very close so that it wouldn't be able to tell which is which. Interestingly, 2 of these people actually managed to fool the performer also, with their fast movements.

All in all, people found beating the computer system a hard task. This appears to be partly because people could not think of ways to move the cups other than those demonstrated in the previous tricks and partly because people had no idea how to go about fooling a computer: "*it's gotta be ... how can you trick a computer? ... I've no idea*". This also relates to the issue of trusting the computer that came up in the In Front of Your Eyes trick. Those who beat the computer, expressed satisfaction in managing to beat it - "*it felt GOOOD [to beat the computer].*"

Most People Don't Mind Being Tricked By a Computer All the audience members said that they enjoyed the tricks. In terms of the questions about disappointment, 3 people (18%) expressed disappointment about the use of the computer in the tricks - one said that on being shown that the computer was doing the first trick, their initial reaction was "*I don't trust magic - [I felt] slightly cross*". Another expressed disappointment that magic tricks are not real magic - "*I'm a bit disappointed that there's no such thing as magic ... I realised that all those magic tricks have nothing behind them*", which was slightly unexpected. Finally, one more technical person expressed disappointment with the way that the computer guessed which cup the bead is under correctly during the In Front of Your Eyes trick, "*I thought that ball must be special, RFID, something, I don't know [but] you just pushed a button!*". As for why the other participants did not feel disappointed by the tricks, "*[it's not cheating], it's part of the trick I think*", "*Not really, I knew it was a trick anyway*", "*everybody knows that there will be a trick*". One participant thought the use of a computer to do the tricks made it better - "*amazing technology ... it was quite clever, a lot of tricks, when you find them out they're quite lame really*".

The framing of the tricks as magic fitting them into a genre which was already well known to all participants was clearly helpful as people did not mind being tricked within that familiar frame. As well as this, the way in which the reveals were integrated into the tricks also seemed to be a good strategy.

Expectations and a Misdirection Failure People's expectations proved key to the misdirections in these tricks. As mentioned above, the first trick was successful in part because on seeing the plain-looking wooden and cloth setup, none of the people expected that the trick involved a computer, or any kind of electronic technology. The In Front of Your Eyes trick was also supported by expectations. At this point, people are expecting the computer to do the clever part, rather than the performer - "*you tell people how you play first time, you use the computer ... I don't realise you actually play the trick*". The expectation is also helped by blurring the boundaries between the first trick, the reveal, and the second trick, with the second trick being introduced as a demonstration of the technology used in the first trick. This means that people are more likely to keep their eyes on the screen during that trick.

The one time when an audience member saw through the trick, video analysis reveals that what went wrong was that the performer was too hurried in this initial section, and attempted to perform the switch before the audience member had fully understood the link between the computer screen and the events occurring on the table - this meant that they were still watching his hand movements as

Believed Cover Story	Developer	Non-Developer
y	80%	29%
n	20%	71%

Table 5.4: Effect of Software Development Experience on Believing the Cover Story

well as the screen. The key to working that trick was to perform the switch of the bead late enough so that the person is just looking down and tracking the cup on the screen, but early enough so that they have not yet realised that a trick is being played. How the cup was moved was also important here. By doing a few big moves with the cup being followed to start with, the performer could draw the audience member's eye to the screen and make the tracking really obvious. In the failed trick, the performer did the switch too early in this sequence. This is again about expectations, this trick can only succeed once the expectation that the movement of the cup will be reflected accurately by the screen has been lodged firmly into the mind of the audience member.

Effect of Knowledge on Trickery There was some statistical evidence of an effect of technical knowledge on people's ability to trick the computer, and likelihood to be tricked by the magic tricks. 10 of the participants had some software development experience. Of these, 8 (80%) believed that the Cups by Sound trick was done by sound alone, as described in the cover story. In contrast, of 7 users with no development experience, only 2 (29%) believed the cover story (Table 5.4). This result is only 94% significant (P=0.058 Fisher's Exact Test) but is directly opposite to the initial hypothesis that technical knowledge may make it harder to trick people using computers. This may in part be because they were more able to come up with explanations for how it might be possible - "*You counted the number of times I moved it*", "*You could hear the bead moving in the cup*".

When fooling the computer, technical participants did seem to have a slight edge, with a larger percentage of non-developer users failing to fool the computer completely (42% vs 20%). The 2 participants who fooled the computer on the 1st try were both developers (Table 5.5). From observing this, developers seemed more likely to work out that the system was only tracking the cups and not tracking the bead at all, which made it much easier to trick the computer.

Knowledge of magic did not have any significant effect on the suggestibility of

Tries	Developer	Non-Developer
1	20%	0%
2	10%	29%
3	50%	29%
Gave Up	20%	42%

Table 5.5: Effect of Software Development Experience on Fooling the Computer

Believed Cover Story	Seen Live Magic	Not Seen Live Magic
Yes	55%	63%
No	45%	37%

Table 5.6: Effect of Magical Knowledge on Suggestibility

Believed Cover Story	Performed Magic	Not Performed Magic
Yes	45%	67%
No	55%	33%

Table 5.7: Effect of Magical Experience on Suggestibility

the participants - those who had seen magic live showed no differences in levels of belief compared to those who had only seen magic on television. Similarly, there were no significant differences between those who had performed tricks themselves and those who hadn't (Tables 5.6 & 5.7).

Gender Effects In some of the older magic literature eg. [129](p492), women are described as being more suggestible than men. This was not the case in this experiment - of the men doing the experiment 78% believed the cover story in the first trick, compared to 38% of women. This suggests that the men in this experiment may be slightly more suggestible, or at least that the women were not any more suggestible, although this may be just due to random variation (P=0.117 Fisher's Exact Test), there is certainly no evidence to support the claim that women are more suggestible than men (P=0.987, Fisher's Exact 1-tailed) (Table 5.8). It is also worth noting that a larger percentage of the developers were male (80% vs 38%), which makes it hard to separate the effects of gender and technical skills. A larger and more randomly sampled study would be needed to analyse these two issues in more depth.

Believed Cover Story	**Male**	**Female**
Yes	78%	38%
No	22%	62%

Table 5.8: Possible Effect of Gender on Belief on Suggestibility

5.7 Conclusions

This development has demonstrated the usefulness and practicality of the deception framework in the design of a complete performance. As well as this it has highlighted several important factors for the successful use of systems designed with this framework.

Key to this work is the framing of a situation and the expectations created by this framing. Whilst systems themselves and direct interactions with them are often studied in detail, this performance demonstrated that what is occurring around the interaction may be equally important.

The framing includes both physical and temporal elements. The physical framing of the cabinet design is used to make the cover story in the Cups By Sound trick more plausible. Temporal framing involves considering things that happen before and after the situation. For example the In Front of Your Eyes trick is reinforced by breaking expectations created in the first trick.

Another important role of the framing is the creation of trust in the audience. By framing this performance as a magical trick a situation is created so that the audience members do not feel hurt that they are being tricked, instead, as described by Jastrow [71] '*he knows that he is being deceived by skill and adroitness, and rather enjoys it the more, the more he is deceived*'. This allowed the exploration of the potentially difficult situation of deceiving another person, and later revealing that deception, without causing any harm or anger.

A final use of framing demonstrated by this research is that things such as the deliberately low technology wooden and cloth table setup and the performer pretending that he is listening to sounds can be used even to misdirect people who have in depth knowledge of the technology.

As demonstrated here, framing is a key issue that must be considered in the use of the framework. This will be considered further in Chapter 7.

Chapter 6

Rock

Rock is an interactive art installation which presents a rock in a cage as if it were a living creature. When the rock is touched, it makes animal-like noises. The noises respond to the way in which the rock is touched, and the pressure of the hand on the rock. By subverting the norms of how a rock responds, Rock creates an intriguing experience which engages, encourages, endears and sometimes scares those interacting with the rock.

This installation explores suspension of disbelief within an interface. In particular, it is designed and presented to encourage multiple people to interact with and around the rock. This creates an interesting situation, where the suspension of disbelief is reinforced by the group experience.

Rock uses a novel computer vision algorithm developed for this project in order to sense pressure using just the view from a web cam in the top of the cage. A large range of different sounds are created using a custom sound engine, in order to make the sounds believable.

This project has been presented to the public on several occasions, and has now been seen by approximately 150 people. This has allowed for the observation of how people interact with the rock, and how this is affected by the group dynamic around the installation.

6.1 Why?

Rock explores the power of knowing deception, where a person knows that they are being tricked, but chooses to suspend their disbelief and act as if they believe the reality being presented. This is focused tightly upon a single deception, that of the pet rock. This installation is designed to have multiple people interacting with it, which allows us to explore the way in which group social factors have an effect on the success of the deception. In particular, it uses a strategy in which a performer initially introduces the installation to early viewers, but once a few viewers have seen the installation, they are left to their own devices and these viewers introduce the installation to others. This allows for the inital interpretation provided by the performer (and also by the framing of the installation) to be expanded and mutated, as each person's unique interpretation of the rock becomes part of the overall story.

This project also demonstrates several elements which are designed to encourage the basic idea of the rock as being a pet, rather than an interface to a computer system, in order to avoid the essentially task-based interpretation that would imply. This is a careful balance between pushing a particular interpretation and allowing the audience to add richness to the installation by adding their own interpretations over and above the initial framing. The use of stone, which is not a material typically used in computer interfaces, is particularly effective in avoiding the interpretation of the rock as being a user interface to a computer.

In this chapter, the design and development of the rock is described, followed by a description of the ways in which people reacted to it when it was shown and how these can inform designers wishing to create this kind of suspension of disbelief.

6.2 Related Work

The concept of a pretend pet is quite an old one. Stuffed toy bears, dogs, cats, octopuses and many other creatures have been amongst favourite children's toys since at least the beginning of the 20th Century. At one point in the 1970s, there was even a brief craze for pet rocks, which were basically a stone, with a pair of toy eyes stuck onto them. A few of these toys had simple mechanisms inside, which allowed them for example to bark, or to walk along, but basically they were static objects, made into something more exciting by the power of a child's imagination.

In 1995, computer game company Ubisoft released the Petz [162] series of games; computer games which simulated an animal. These virtual pets could be fed, taught tricks and in later versions could even be bred together to create virtual offspring. Soon after this in 1996, the Tamagotchi [151] was launched, a

Pictures ©http://www.flickr.com/people/the_pink_princess/, Joshua Wickerham, Jörg Kantel

Figure 6.1: A Tamagotchi, a Furby, and a Sony AIBO

virtual pet which lived in a small digital watch-like electronic toy. This was soon followed by the Furby [161], a small furry creature, which had similar needs to be fed and played with, but interacted by making noises, waggling its ears and rolling its eyes and used sensors to detect touch and light. Another virtual pet was the Sony AIBO [159] dog, which could walk around and even had a cycle of development of its personality, from puppy to fully mature. These various virtual pets are shown in Figure 6.1 .

In many ways, commercial virtual pets are a step backward from the simplicity of static soft toys. They are constrained in many ways by the particular interactions designed into them, unlike soft toys. A teddy bear for example can be anything from a small friendly creature that lives on honey [153], to a space going super hero [163]. Rock tries to avoid these constraints using a design strategy inspired by HCI and design work which creates objects without a fixed purpose. For example the History Tablecloth [52] is a tablecloth which lights up when objects are placed on it. This creates an intriguing interaction without a particular meaning, upon which the user may apply their own interpretation rather than being designed to be interpreted in a particular way [116].

Rock is designed to create interpersonal interactions between people around the installation; in this aspect it is similar to the Paro robot seal which was also designed to encourage social interaction [91]. It is also designed to be socially discoverable, where one person using the interface may attract and inform others about it. This kind of interaction is similar to that observed by vom Lehn et al. with simple computer vision interfaces in museum settings [130].

Another project designed to create interesting responses without suggesting a particular interpretation is the Meatbook [84]. This is an artistic installation which uses a tangible interface made out of meat with embedded electronics and

Figure 6.2: The Rock in its Cage

motors. Meatbook is designed to create interestingly visceral responses in viewers. Like Rock, this is an object which due to the choice of material is surprising in its interactivity.

One other related art project is Pony, an interactive artwork by Tim Lewis [148]. This is a strange robot creature with arms and hands for legs, and a neck made out of a single arm with a pointing hand on the end, pulling a two wheeled pony cart. It uses ultrasound sensors to respond to nearby movement, with the neck moving, and the legs walking around. This creates an eerie creature which comes up to people and gently reaches out to touch them, and is surprisingly anthropomorphic yet very much unclear in its intentions.

6.3 Technology

The rock is presented in a cage, with a small door allowing people to reach in and interact with the rock. At the top of the cage, a camera is clipped onto a bird perch, pointing down at the top of the rock, to get a good view of people interacting with the rock. Speakers and a computer running the rock software are hidden underneath the table (Figure C.7). The table is covered with a black cloth to keep all the technology hidden from view. The cage makes sure the rock is always in view of the camera, and means that people may only interact with the rock with one hand at a time, which makes the tracking algorithms work more reliably.

The software in the Rock is split into two parts. Firstly, sensing of hands (and potentially other objects) interacting with the rock. Secondly, a sound engine, which creates responses to these interactions.

6.3.1 Hand and Pressure Sensing Algorithms

The rock detects touch using a custom pressure sensitive finger tracking algorithm which works by looking at the fingernail from above using the video camera in the top of the cage, detecting changes in colour that result from blood moving under the nail due to pressure. For full mathematical details of the algorithm see the paper presented at Pervasive 2008 [90] (included as Appendix C). Interestingly, there is one other related project developing fingernail-based pressure sensing which uses a more conventional pattern recognition algorithm, designed to spot particular actions, rather than the relative pressure-based algorithm here [125]. The algorithm described here is also novel in that it does not require calibration, in contrast to the previous approach, which requires calibration with an industrial robot to apply known amounts of pressure to a persons fingertip.

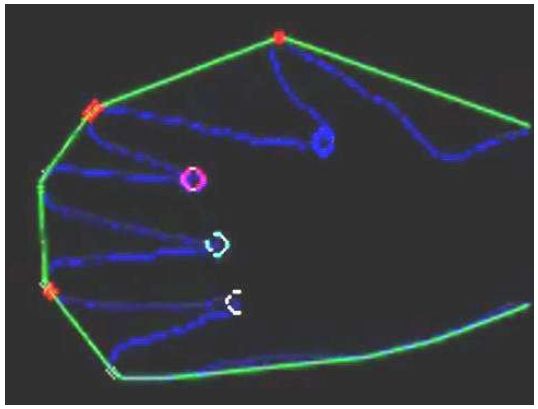
This algorithm gives two outputs, the position of the hands and the amount of pressure which the fingers are applying. One particularly interesting element of these is that the hand position is clear to onlookers, whereas how hard the fingers are being pushed is only really visible to the person pushing. These two levels of visibility are used in the design of the rock interaction.

The 4 stages used in the algorithm are described below.

Background Segmentation First, a simple colour-based background segmentation is performed. This segmentation uses a known background colour (as the colours of the rock and the cage are relatively uniform) rather than any kind of skin detection. This is deliberate so as to make the rock respond to other objects placed in the cage as well as to hands. It is also somewhat more robust to lighting changes than skin colour detection, as it only needs mask out a small range of cage background colours so can be set quite loosely without causing problems. This gives a bitmap showing a silhouette of the hand.

Hand Detection The outline shape of the hand is detected using the OpenCV library contour detection functions. The convex hull of the hand outline is calculated, along with the convexity defects, areas where the outline goes away from the convex hull. This is used in order to detect the position of the hand and fingers, and to detect the shape of the hand. One key output from this detection is the position of the ends of the fingertips. An example of this output can be seen in Figure 6.3(a).

(a)Detecting the hand shape & fingers



(b)Detecting the fingernail and sensing pressure.

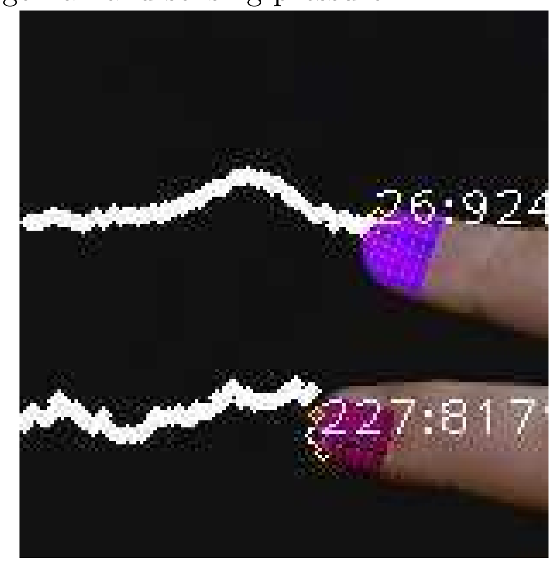


Figure 6.3: Pressure Tracking Stages

Fingertip Detection Using the endpoints of the fingertips detected in the previous stage, the rough area of the fingertip is detected using a simple estimate of the relative length of fingertips compared to the width, that fingertips are approximately 150% as long as they are wide. Due to the pressure sensing algorithm used this estimate does not have to accurately detect a nail, it just has to consistently detect roughly the same amount of a finger as being the fingertip. In figure 6.3(b) the detected fingertips are shown highlighted.

Pressure Sensing The graphs next to the fingers in Figure 6.3(b) show the pressure detected at each finger. Pressure is sensed by looking at the fingertip. This sensing uses the fact that fingernails look different under differing amounts of pressure. When the fingernail is pressed down on a surface the blood under the nail is concentrated into one area (Figure 6.4). This means that the fingernail goes from being a roughly uniform colour, to having two distinct areas each of which are different colours. A simple measure of this is to measure the variance of the colour of the pixels that make up the fingertip (this uses the hue variance, calculated using a vector conversion. See Appendix C for more details). This gives a measure of pressure which changes relatively smoothly as the finger is pressed down harder. This measurement is relatively constant under rotation and at different distances from the camera. In this system, the fingertip values are normalised based on the range of of recently observed values, which allows it to give a useful pressure reading for different individuals and lighting conditions.

Figure 6.4: A fingertip with different amounts of pressure

An interesting side effect of this simple pressure sensing method and the hand detection algorithm used, is that the pressure sensing actually works for grasping the rock as well as simple pushing. In this situation, it works because the knuckles and any visible part of the forefinger are detected as 'fingertips'. As the rock is grasped harder, the balance between knuckle and (differently coloured) forefinger in the detected fingertip alters. This means the pressure algorithm gives a response to grasping the rock - which is a common interaction mode (Figure 6.5). This gives this algorithm a significant advantage over algorithms which attempt to detect the exact shape of the fingernail on the finger, or use other shape-based estimators of hand pressure.

In the case that a person touches the rock with an object, the hand detection and thus pressure sensing fails to work, and will typically give no values at all, as it cannot detect the finger silhouettes consistently. However, in this case the outline is still used to output a hand position value to the sound system, so that the rock will respond to these touches, although it has a less nuanced response.

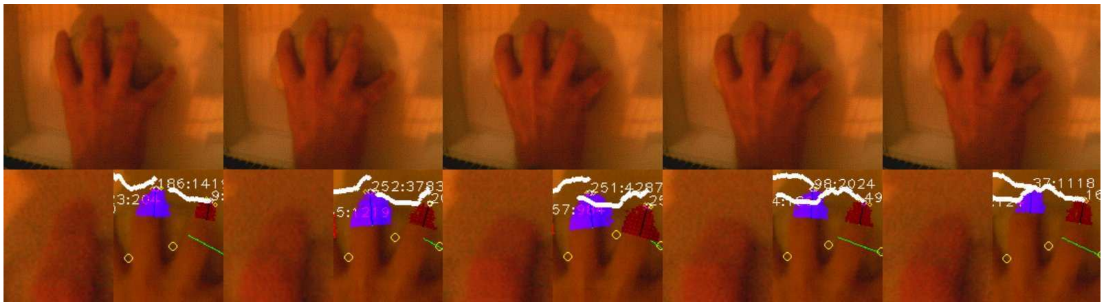


Figure 6.5: Grasping the Rock

6.3.2 Sound System and Personality

The input from the hand tracking and pressure sensing is converted into two metrics, ‘Excitement’ and ‘Fear’. ‘Excitement’ is a measure of how excited the rock is. This is based on the pressure sensed by the rock. ‘Fear’ is calculated based on how fast the hand (or other object) moves towards the rock. This means that if people grab at the rock quickly, it gets scared, if they poke it hard it gets excited, and conversely, if they touch it gently and slowly, it will be more relaxed and not frightened. The values for the two metrics are constrained so that the rock does not always instantly change temperament, for example it can get scared very quickly, but will take some time to relax after it has been given a scare.

These metrics are used to pick sounds from a large library of sounds, which are basically animalistic noises. These are categorised as to where they fit on the excitement/fear metric, and emitted depending on the current level of the metrics. For example, when the rock is scared, but not very agitated, a low pitched growl may be made. The sounds output are also altered in pitch during output based on the agitation level, which gives a more nuanced output which alters depending on how the rock is being handled. There is also a heartbeat sound, which alters in speed and volume depending on how scared the rock currently is. This heartbeat also serves as an attract sound, to draw attention to the rock, and to suggest interaction with it, which is particularly useful in an artistic setting, as it avoids people seeing the rock as simply an artwork to be passively observed.

6.4 Performance & Installation Design

Perhaps the most important part of the design of the Rock installation was finding a pleasingly shaped rock that was interesting both to look at and to feel. After trying a surprisingly large number of rocks a suitable rock was found in a search of the author's garden. Once the rock was found, the installation was created using it as the focus for the design. Several prototypes were created, with variations of the sound output system and the physical framing, in an iterative process with some small scale user testing. This resulted in the final design described here.

The design of the Rock installation is very much about framing the rock and encouraging people to disregard the technology involved. The most important part of this is the use of the cage. This forces all interaction with the rock to occur in the view of the camera. This avoids situations where the tracking breaks down when it cannot see someone's hand, and preserves the illusion of the rock as an interactive object itself rather than as part of an interactive computer system. Hiding the computer and the speakers beneath the rock's table also aids this illusion. It also avoids people taking the rock away, which again would have broken the illusion.

As well as providing the physical framing of the interaction, the cage also has a social meaning, in implying that whatever is inside it is a pet: an animal kept in a cage. A further aid to this interpretation was given by attaching a zoo-like text label to the cage, with a fake latin name and instructions on feeding. This also resonates with the Pet Rock craze of the 1970s, where stones with eye stickers stuck on them were sold as joke pets and became massively popular for about 6 months. Interestingly, this also suggested a particular interpretation of the camera by some participants: that it was a CCTV camera used to monitor the pet.

The mappings between input and sound in the rock were designed to work on two levels, based on the observations made in early testing of the pressure sensing algorithm - that the movement of the hand was clear to onlookers, but the pressure of fingertips was not. The idea behind this was that at first the Rock would be simply an intriguing object. Then, when people saw someone else interacting with it, it could be seen that touching it fast or slow made a difference to how it reacted. Finally, when they got to interact with it themselves, they could see a more subtle response to finger pressure and grasping. This process of increasing knowledge was designed explicitly to create a balance between a sense of mystery and some level of meaningful interaction with the Rock, with the hope that the initial mystery would reinforce people's suspension of disbelief and they would interact with the Rock as if it were a real animal.

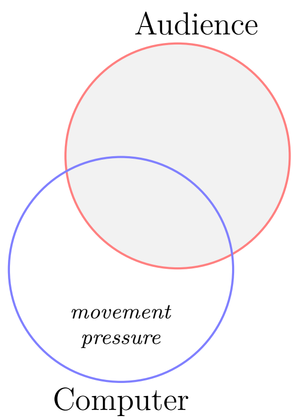
Key to the meaningful nature of the initial interaction and the core personality of the rock was the categorisation of the sounds and the choice of the variables for the response. The choice of variables was made by playing with the system with the basic tracking working and exploring what effect various gestures such as grasping, grabbing, poking etc. had on the raw tracking outputs. The two metrics of fear and excitement were chosen based on this exploration as being evocative of the way a pet animal such as a cat responds to human contact, recoiling from being grabbed at quickly, but enjoying being stroked firmly.

The sounds were chosen by first sitting with a microphone and making a very large number of animal sounding noises (approximately 45 minutes of audio). A subset of these sounds were then chosen as 'good' sounds and each one was categorised against the two metrics of excitement and fear. Clearly this is a very subjective process as the actual meaning of any particular sound is not clearly defined, so sounds may be interpreted differently by different people, but it does give a basic, roughly consistent personality to the Rock.

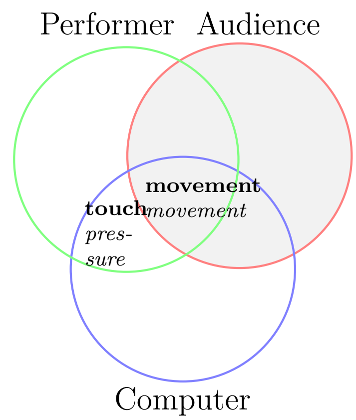
In this installation, there is not a fixed performer, rather, as described by Reeves et al.'s work on public interaction [109], there is at any one time a person directly interacting with the interface, who can be seen as taking a performance-like role and other people who are looking on to see this person's 'performance'. There are two sets of actions relevant to the performance - hand **movement**, and finger **touch**. These are mapped using respectively hand *movement* mapping, and finger *pressure* mapping.

Rock is designed so that people will change from onlooker to performer roles and back again over time. It is also designed so that people go through a trajectory of increasing knowledge about the rock. The stages are shown in Figure 6.6, with the role that the person is taking on shaded grey in each diagram. They go from not knowing what it does at all (Figure 6.6(a)), to knowing that moving a hand towards it makes it respond (Figure 6.6(b)), to knowing how changing the way it is held makes it respond by actually experiencing it (Figure 6.6(c)). Finally they watch it again, with the added knowledge the finger pressure alters the way it responds, but still unable to actually see finger pressure interactions (Figure 6.6(d)).

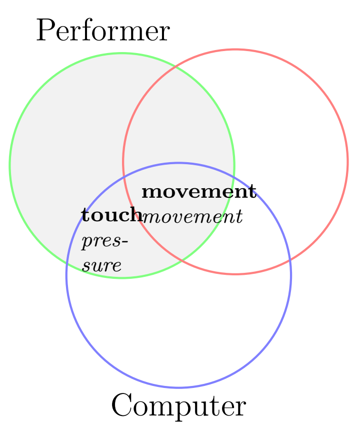
What this design is trying to support is the overall illusion of the rock as a



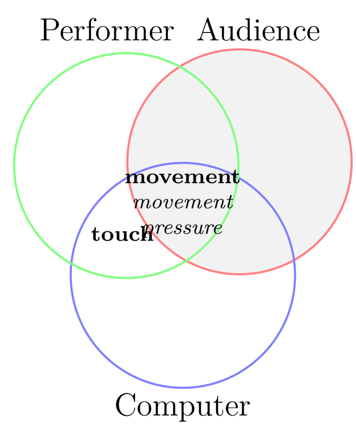
(a)Seeing Rock on its own



(b)Watching another person interact

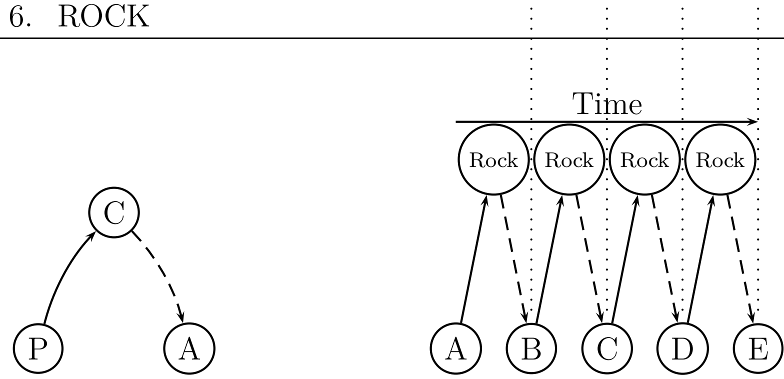


(c)Playing with Rock



(d)Watching Rock again

Figure 6.6: Trajectory of Understandings in the Rock



(a)The initial rock deception (b)Passing on the deception over time

Figure 6.7: The Illusion of the Rock as a Creature

creature. This can be described as a Story Telling deception (Figure 6.7(a)), created by the author as initial performer, with audience members, who know that a rock cannot develop consciousness, suspending disbelief to interact with the rock. However, as well as this initial designed illusion, the aim of Rock is that people who have seen it already themselves support the deception for another person - leading to the situation shown in Figure 6.7(b), where the deception is passed on from person to person over time. A further intention of this style of interaction where knowledge is constantly being passed on, is to allow space for people to interpret the rock in their own way, and to pass on their interpretations to those who see it later. This is intended to create a range of ever changing interpretations of the rock, rather than just the simple high level framing of it as some kind of creature, and thus to give a depth to the rock beyond that provided by the object alone.

6.5 Public Interaction with the Rock

The rock was taken to the Reactor Digital Live Art Conference, where it was exposed to approximately 120 people. This allowed the idea of creating a suspension of disbelief to be tested with a large group of people.

The rock was set up in a quiet corner of a corridor, which was chosen as it was between two rooms where people would be passing through during the sessions and breaks in the day. This positioning allowed for the noises made by the rock to be heard, which ensured that most people noticed the rock. The volume of the heartbeat sound was set loud enough to catch the interest of passers by. A typed note was put on the cage of the rock, based on a zoological description of an

Figure 6.8: Rock in its first public guise

animal complete with false Latin name. This note suggested that the rock enjoys being touched, but was rather shy, and was aimed to encourage viewers to interact with the rock and to see it as having a personality (see Figure 6.8). Rock was also shown at various computer lab open days and presentations, which allowed further observation.

A particular focus of interest in running the installation, was how the group setting would affect the effect of suspension of disbelief, and whether this would reinforce, or make things worse. In practice, it seemed extremely effective and engaging - the Rock was awarded best paper prize at the conference. The following sections describe in detail the way in which the rock worked in public.

Group Effects and Critical Mass During initial observation of people looking at the rock it became clear that a lot of the people there did not feel confident to interact with an object that was being presented as art, despite the attached notices suggesting touching the rock. Some people even worried that the whole thing was an elaborate trap to catch people with an electric shock or similar. Because of this, rather than purely observe the rock in a hands off manner, the rock was presented as a pet, with the author acting in the role of 'owner' encouraging viewers to break through this barrier and reassuring people that nothing bad would happen to them. The owner had detailed knowledge of the inner working of the rock and practice in interacting with it, so was able to attempt to make it calm down and sound happy if it got over-excited.

After about 30 minutes of the owner introducing the rock, something interesting began to occur; people who had previously played with the rock came back, bringing other people to see it. These people then assumed the role of the rock expert, meaning that the owner was able to let people interact with the rock with no reassurance or demonstration required. In this large group situation, with the initial catalyst of the owner, people really did go through a trajectory of understanding somewhat similar to that shown in Figure 6.6. One basic example of this is the mapping of the hand movement. Typically a person playing with the rock for the first time without having seen it played with before would reach very quickly into the cage and touch the rock, causing it to make snarling noises and them to pull their hand out quickly. People with greater knowledge in the care of rocks would then suggest or demonstrate that they should touch it more gently and slowly, which then led to them discovering how to make the rock respond more favourably to them.

Other People Encouraged Suspension of Disbelief As well as providing a useful way of introducing the installation, having a large number of people around did seem to encourage suspension of disbelief. Even those who were initially only interested in how the system worked were drawn into the idea of the rock being a creature by others viewing the rock. Some people even encouraged others to not look at the technology, telling them that the technology wasn't the point of it.

Development of Multiple Interpretations Most people were unable to see how the rock worked, as they could not see any computers, screens or other obvious interface devices. This, along with the framing helped guide almost all viewers into interpreting it as a creature in some way or other. The way in which people interpreted the various sounds varied a lot, with one particular sound being described by onlookers as everything from a happy purring, to a scary growl. The ambiguity of this simple output, appeared to lead people to create a much more rich and varied set of stories about their interaction than had been initially designed. This use of ambiguity, and suppression of obvious interpretations was originally inspired by HCI work on ambiguity and multiple interpretations [50, 116]. In this case, these strategies worked extremely effectively, creating a design which deliberately went beyond what the designer could envisage, by allowing the users to create their own level of interpretation, creating a personality for the rock that went far beyond the original idea.

People Really did Suspend Disbelief When people interacted with the rock, the idea of it having a personality was definitely influencing their behaviour. When people did not initially understand the way the rock was responding to them, they did not blame this on the rock, as they had seen it respond nicely to others, they described it as the rock not knowing or liking them yet, and other anthropomorphic terms. This effect was encouraged by the owner, who avoided referring to the technology involved in the rock, telling people that they could read the paper if they wanted to know more about the workings of the rock. A measure of the success of this approach was that several people took photographs of themselves with the rock. One person even approached the owner at the end of the day and asked if they could take the rock home to live with them.

Rationalisation of Unexpected Events One of the most interesting aspects of these user-generated interpretations was the way in which they are extended in the event of an unexpected event. For example, the lack of a known set of gestures for interacting with the rock meant that some users performed actions such as tapping the rock or turning it over, which caused it to react relatively randomly (although it consistently did provide some reaction to these gestures). Rationalisations were made for these responses, such as that ‘he doesn’t like being upside down’, or ‘he likes having his tummy tickled’ (it is interesting to note that most viewers seemed to think that the rock was male). Rationalisations were also made when the rock responded to the light turning off in the area that the rock was set up by making random noises. It was described as being scared of the dark, and also by a different person as going to sleep and snoring.

Randi, arguing against false psychics describes similar rationalisations being observed by those believing in such phenomena [106], and how they persisted even when he told the believer that he was performing a trick. In this case, it appears that they even hold when people are completely clear that it is just a technological trick - people are ascribing significantly more artificial intelligence to the technology than it actually has.

Unobtrusive Technology The technology used in the rock was quite unobtrusive. Most people either didn't notice the camera, or thought that it was simply a prop of some kind. Those who did notice it sometimes played a little to see what the camera was up to, but it did not appear to destroy the suspension of disbelief, they typically went back to interacting with the rock as normal. This demonstrates how in a successful suspension of disbelief, even if people do see the technology, as long as it is made easy for them to choose not to pay attention to it, the visibility of the technology is not a problem.

Interestingly, the sound coming from under the table was not a problem either, and no-one looked under the table. A few people even lifted up the rock and turned it around in order to check there were no speakers in it. This relies on the fact that people cannot reliably localise the source of a sound, a deception that has been used by magicians throughout history [129].

Even the Developer Suspends Disbelief During the design of the rock, I spent several weeks living with and testing various versions of the rock. During the later development of the rock, as the personality was programmed, I found myself increasingly becoming attached to the rock. Despite knowing exactly how the sensing and audio processing was working, the idea of the rock as having a personality became increasingly believable, especially at the point where the animalistic noises (which I made) were added in. Family members & friends who saw the rock and had its working explained to them also exhibited this ability to suspend disbelief and to conceptualise the rock as an animal, despite knowing that it is simply a well disguised computer program. This experience of suspended disbelief helped very much when posing as the owner of the rock in public.

6.6 Conclusions

In many performance and artistic situations the creation of an 100% realistic illusion is not the aim. Rather, illusions are created that only hold up with the collaboration of the audience. These use the 'story telling', suspension of disbelief illusion. For example, theatre only works if the audience decide to consider the actions on stage to be real - this suspension of disbelief is vital to the success of much live performance. Even in stage magic, people are often tricked only because they choose to be tricked [99]. This understanding is key to the use of suspension of disbelief to tell stories.

This chapter has described the development and display of Rock, an interactive artwork based on a pet rock living in a cage. This system creates a suspension of disbelief in the people interacting with the rock, who accept that it is a creature with a personality. Showing this to large numbers of people in public enabled observation of the way in which groups of people have effects on the suspension of disbelief.

In terms of how it can inform the deception framework, this work identifies two interesting elements of the framing of a situation which have not been considered in the previous chapters. Firstly, the effects that groups have on deceptions. Secondly, the use of a trajectory through an experience to add depth to deceptions. These two effects are summarised below and will be discussed further in the revisiting of the framework in the next chapter.

Group Effects One key observation from Rock, is that groups of people can reinforce suspension of disbelief. This effect occurs because as people interact with the installation and each other, they choose to suspend their disbelief, and indeed there is a social pressure to do so. This is similar to the observations by magicians that it is easier to trick multiple people than one [111], and that in large group situations even if one person spots the trick, they often will not say anything, for fear of spoiling it [99]. It is in the interests of people's entertainment that they help maintain the story.

This group interaction can take the load of creating suspension of disbelief off the performer. It also has the benefit of making the illusion even better than imagined by a designer, by fleshing it out with details created by the audience members themselves. This ceding of power to the audience is also useful in situations where the user is interacting in ways not envisaged by the designer, or when the technology fails for some other reason. Similar effects were also noted in Gaver et al.'s History Tablecloth [52]: when the technology was not doing something obvious, it was an interesting point of discussion, rather than being always seen as it 'stopping working'. One important thing in both the History Tablecloth, and Rock, is that in the event of an unexpected input, they either reacted incorrectly, or did not detect the input, but would still limit their response to things consistent with their normal working, rather than producing error responses. This allows people to build these failure events into their stories about the system, and again helps maintain suspension of disbelief.

The final point about group illusions, is that creating illusions in a performance or installation shown in a group setting is actually easier than in an individual settings. This is very much due to the situation in which this kind of work is presented. The default position of an audience is to want to be entertained - in effect to help the performer, and as long as a majority of audience members are on the side of the performer, social pressure will keep the audience believing. One example is that in this installation, despite the camera being visible, and the computer, screen and speakers being beneath the table, not even one audience member tried to look behind the cloth covering the hidden technology.

Framing and Trajectories One other factor that was clearly important is the initial framing of a situation. The experience of the rock shows that (quite possibly with justification) people may be afraid to interact with art works, particularly when there is a risk of looking like a fool in front of others. In situations like this, where explicit instructions are not suitable, it may be better to have a performer to demonstrate the trustworthy nature of the work and encourage people to use it. However, this installation also demonstrated that once a few people have trust in a work, this can spread to others, reducing the need for this performer. As well as the initial framing the fact that a trajectory through the experience existed was important for the performance to work, the designed trajectory of gaining knowledge about the rock meant that people would have an incentive to interact with the rock once they had seen others use it. Interestingly, people also were keen to pass on their newly found (and created) knowledge about the rock.

Chapter 7

Discussion: Revisiting Deception

Previous chapters have discussed the use of deception within three very different performance situations: a staged performance, a close up individual performance, and an artistic installation. This has both demonstrated the usefulness of deceptive techniques in computer augmented performance and has allowed for the exploration of these techniques in real world settings. This chapter presents a general discussion of the deception framework and revisits the 7 types of deception described in the initial framework chapter, informed by the complexities observed when using deceptions in performance settings.

As well as this more detailed explanation of the performance use of deception, this chapter discusses the possibility of extending the definition of performance in several ways; to performances with more complex roles, to remote or asynchronous performance, and to encompass activities that occur in other areas of HCI .

Finally, the question of the ethics of deception is discussed. Whilst deceptions may be useful to achieve outcomes desired by designers, care has to be taken in their use, particularly if a situation is being created where people do not know that they are being deceived, or where a clear framing of the situation is not provided. This chapter discusses the positions taken by other professionals who use deception and what may be appropriate for performers using interactive technology.

7.1 Revisiting the Framework

This section revisits the original framework in the light of the findings of the three case studies, and presents a more detailed explanation of the complexities inherent in the application of these techniques.

7.1.1 Framing of Deception

How interaction is framed, or what goes on around the interaction and to explain the interaction, has been shown to have a significant effect on several aspects of HCI. The most well explored aspect of framing is the effect on website design. For example, it has been shown that the purpose for which people are using a website has a significant effect on how attractive it is seen as by them [3, 58]. Also, how questions are framed has been shown to have a big impact on people's responses, for example if people are asked to opt-in or opt-out of the use of their personal data [80]. Similarly, when the same statistic is used to give people information about a website, whether it is framed in a negative or positive way (e.g. 10% of people found it hard to use or 90% of people found it easy to use) has a significant effect on their rating of the website quality [58].

This section discusses the use of framing effects when planning deception based interfaces. In many ways, this is the same as when planning a non-augmented performance. However, the factor of relevance to this thesis is the level of integration that is required between the external framing of the deception and the design of the system itself, with physical framing having an effect on the software design, and vice versa.

Framing Deceptions within Other Deceptions It is clear from the Cup Game case study that framing effects are important to the use of deception, and to ensuring that people are not annoyed or disappointed by deceptions. Several people said after that performance that they knew they were seeing tricks so they didn't have a problem with being fooled, even though they did not know exactly how they were tricked. In this situation, the hidden deception is being framed within a wider deception, which is that the whole thing is a set of magic tricks, which the performer is clear about. This avoids making people feel stupid, angry or disappointed when the trick is revealed, even though they have been deceived, which is key to the performance - as the magician James Randi says '*I mislead the audience in order to create wonder in their minds, and thus entertainment*' [106], rather than intending any kind of disrespect for the spectators.

Physical Framing The physical setting in which the deception occurs has to be considered when designing the deception. The settings used for the Cup Game and Rock were designed to support the deceptions being used. With the Juggling

Tracker, the deceptions which emerged are in some cases reliant on the physical framing inherent in stage performance.

In the Cup Game the design of the table used to show the cups and the space in which they were shown was central to the framing of the tricks. The table is designed to resemble a stage, with curtain like cloth around the sides and top and strong top lighting onto a clean wooden surface. The booth design in which it is set up also resembles a magician's cabinet, Punch and Judy shows, fortune telling booths and other variety/fairground style performance settings. These strengthen the framing deception, that this is a magic trick, creating a framing which audience members are able to understand, and know how to act within, due to past experience and knowledge of such settings.

In Rock, the physical framing by putting the interface in a cage and making the technological parts of the system unobtrusive, is vital to maintaining the suspension of disbelief. Clearly in a suspension of disbelief deception like Rock this hiding is not so important from a purely practical point of view, as even without seeing the computer someone is not going to truly believe the rock is alive. However, hiding the obvious computer interface elements avoids one of the more obvious undesirable interpretations which is that people see it as just an interface to a computer system and play with the rock whilst looking at the computer screen - which certainly does occur if the screen is visible.

The deceptive control methods in the Juggling Tracker exploit the physical setup of a stage performance in a way similar to that of a magic performer. The fact of the performer being on the stage has itself an important effect of giving control, as the audience know to look at and pay attention to the performer. This makes it easier to control what audience members are paying attention to at any particular time. This physical distance can also make some things less noticeable to the audience, which can be useful if inconspicuous or magic-like control of the system is desired.

Social Framing The Juggling Tracker was primarily framed as a typical piece of stage entertainment. This makes use of the existing social conventions around the performance. This kind of framing is very practical, as audiences, venues and performers are all familiar with the way these situations work. As Nardi [99] describes these expectations mean that the audience collaborates in the framed activity (what is often described as having the audience 'on your side'), and each of the audience members by their collaboration is supporting the performer's efforts to maintain the collaboration of others.

The Cup Game was presented as an individual experience, in a small dark room. This was designed to create a secret, magical atmosphere. This is in part to support the framing as a magic trick, it is also designed to give a feeling that people are having secrets revealed to them, and to encourage people not to tell others the secrets - this was important with participants who turned up for the experiment with friends who waited outside before taking part in the experiment themselves.

In contrast to this introverted atmosphere of secrecy, Rock was set up explicitly to encourage social interaction. It was installed in a high traffic area, where large numbers of people would be passing through, with the cage and the rock visible from approximately 270 degrees. This meant a small crowd of people could all see the rock being interacted with. This framing encourages people to discuss the rock, and to create interpretations of its behaviour between them. This approach removed a lot of control from the performer / designer, but allowed for a richer, socially generated interpretation of the deception that the installation was created around.

7.1.2 Trajectories, Transitions and Reveals

As well as considering what occurs physically and socially to frame deceptions, these case studies have also demonstrated that it is important to consider how the deception is framed in time, and how the deception may alter over the course of an interaction.

Reveals & Transitions As well as the use of different deceptions at different times, it is important to consider how people may discover deceptions during an interaction. Rock and the Cup Game both have elements of discovering more about the deception during the experience. They do this very differently. In Rock, as a person interacts with the rock, they gain more knowledge about the responses of the rock. In this case, the system is revealing itself by a learning process. In the Cup Game, the performer directly reveals each deception to the audience member.

Handling reveals or transitions is an important aspect of performance, in part because this is the point at which someone is potentially going to be made to feel stupid. In magic, this situation is handled very neatly by the social understanding of the situation, which is that the magician is a very skillful and sneaky trickster, rather than that a stupid person is being tricked. The other thing that magicians do in this situation is to not reveal the method behind their deceptions, instead just revealing the fact that the person has been deceived, and leaving explanations up to the audience. Possibly if they revealed how simple most tricks are, the audience would be less happy; in fact one of the audience members in the Cup Game suggested exactly this - "*a lot of tricks, when you find them out they're quite lame really*".

Trajectories The way in which a person goes through an experience is key to how they experience each part of it. This temporal trajectory of the experience may be designed in order to support the deceptions taking place within it.

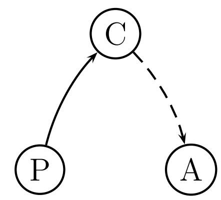
In Rock, there is a loosely designed temporal trajectory, throughout which people learn more about the rock's interactions. This is driven by the different visibility of the possible interactions with the rock and the social interactions about it. The design is based on one person being further through the interaction trajectory than those earlier in the interaction. Benford et al describe this in terms of a combination of trajectories through roles occurring in tandem with a trajectory through the narrative of the installation [9]. In this case each person has their own narrative and contributes to a larger socially constructed narrative.

In the Cup Game, a trajectory of interaction is also occurring. This time it is much clearer, with each deception leading on from the previous one, with reveals of the deception at various points. In part this very explicit trajectory is also used as part of the deception, by using a supposed reveal and demonstration as part of the next trick. This is a classic magical deception, doing the trick at a different time from the time at which the audience member is expecting a trick [99].

Again, as for framing, the relevant thing about trajectories is the way in which their design relates to the system design. For example in Rock, the trajectory design partly relies on particular aspects of the sensing technology used. In the cup game, the trajectory of the performance was designed in tandem with the physical design of the way in which the technology is hidden and revealed.

7.1.3 The 7 Types of Deception

This section evaluates the 7 types of deception from Chapter 3, and considers their potential uses with respect to the framing and trajectory issues described above.



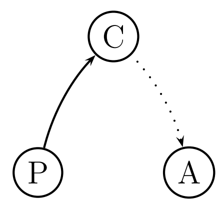
Story Telling As a framing deception, the Story Telling deception is very common. For magicians, the framing of their work as trickery is very important as it prevents people being deceived without their knowledge, or believing they can perform paranormal acts. By actively telling people that deception is going to occur, they are expecting tricks and misconceptions are avoided. In theatre, the stage and theatre are a framing for suspension of disbelief - using well understood pieces of technology, lighting and other conventions that go with them, in order to signal that a fiction is occurring. The Cup Game makes use of this kind of deception as a framing, with the physical design of the setting supporting the deception as described above.

Games, theatre and other forms of storytelling and entertainment commonly make use of technology in order to create suspension of disbelief. This does not rely on the audience believing that something is actually happening, rather to go along with the story. For example, in theatre, when lighting and a fan are used to make it seem like a door goes to the outside, everyone watching knows that this is not the case, but it supports the story being told on stage. The juggling tracker performance described in Chapter 4 is another example of this - the visual and audio augmentation of the performance is used to create comic effects which support the story being told.

The Rock is a Story Telling deception, using the design of the setting and technology to support the audience's suspension of disbelief in the story of the existence of a pet rock.

This kind of deception is likely to be one of the most useful for those designing interactive systems. The big advantage over really deceiving people is that technically it is often far easier to provide a story than to actually fool people. The difficulty in using this type of deception is that it relies entirely on the engagement of the audience, the story being told has to interest them enough for them to play along with it. In many ways this is not a downside, it is simply a reflection of the fact that a performance needs the audience on side to work well - a fact that is well known to performers, and which also has been discussed in relation to other computer interactive media, such as multimedia educational games for children where engaging the audience is key to the success of the educational aims of such software [113].

In a trajectory, the fact that this type of deception reveals information about the mechanics of the deception to the audience means that it is not possible to further reveal the deception. However, as described in the next section, it may be possible for a story telling deception to transition from being the main deception occurring to being used as a frame for further magic trick deceptions.



Magic Trick This kind of deception is used in the tricks in the Cup Game - the person is being tricked by the performer using the computer and they are not aware of it. This means that the tricks can really surprise people, as they do not see what is actually going on. This is part of a wider deception that the audience knows about, yet even though people know that they are taking part in a performance involving being tricked the tricks are still successful. The framing in a suspension-of-disbelief deception is used here to avoid problems with the inner hidden deception, such as people considering it to be paranormal, or people not enjoying being tricked, or feeling stupid at the point that the tricks are revealed as such.

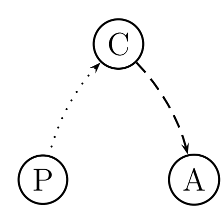
It is also used in the Juggling Tracker when obscuring control movements by embedding them in other performance movements. In this case, it is simply being used to imply that the system is smoothly changing without any extra input from the performer, to create a more seamless performance. In this situation, hidden deceptions are used without any ethical issues, they are not being used to directly affect the audience in any way which has potential to harm them or make them feel cheated.

Using this type of deception as a framing deception within which other deceptions are played may be hard to do without causing distress or harm to participants. For example, many early psychological experiments set up experiments on completely false pretences, such as creating a completely false experimental task, when the experiment itself actually occurred once this task had finished [45], or bringing together a group meeting of local businesses to give them completely false information as part of an experiment [97]. Use of this kind of theatrical technique to make experiments on people was later considered not to be acceptable in many cases due to the potential for harm to participants, and a consideration that it was not appropriate given the experimenter-subject relationship [74]. However, sometimes a performance can be framed with a deception. For example, a staple of magicians who do corporate performances is to appear as a company spokesperson, dressed appropriately and using PowerPoint and similar presenter aids, then surprising people by bringing in magic tricks and a show. In this situation care is taken to avoid making the audience feel exploited by the deception.

Magic Trick deceptions are in some ways harder to create than making use of suspension of disbelief, as everything has to be performed to a believable level. However, there are a wide range of techniques which can be employed for performing this kind of deception, such as misdirections which take people's attention off a trick, and playing with the time of the trick to perform the mechanics of the trick when people think you are actually just doing something mundane. One key advantage of these deceptions is that they can be used to cause a sense of wonder, where people are amazed by what has happened. This is in contrast to suspension of disbelief deceptions, where people can see how things are done.

In a trajectory, it is often the case that magic trick deception may be revealed, at which point it may become story telling deception. One way that magicians use this is to pretend to reveal one trick whilst actually performing the next trick. In this case, the trajectory starts with a magic trick deception (the first trick), which then appears to become a story telling deception (revealing the first trick) which itself is used to frame the magic trick deception of the second trick.

Stupid Computer In this, the computer is being tricked, but this is not fooling the audience member. This is what was occurring in the Beat the Computer trick in the Cup Game when a participant managed to trick the computer causing it to report an incorrect cup to the performer, but it was clear to the performer (who is effectively in an audience role in that part of the performance) what had actually happened from watching them do the trick.

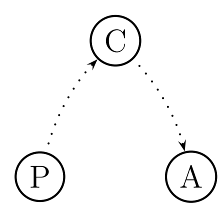


In the Cup Game performance, this kind of trickery was presented as a challenge for people to try. It was also presented afterwards as a demonstration of how stupid the computer system really was - which had some comedic effect, and also helps people learn about the limitations of such systems.

Other than the use as a challenge, or for comedic effect, the other use of this type of deception could be simply as a storytelling strategy. By performing using a computer tricked into thinking something is happening that is not really, it may be possible to create interesting effects that support a story being told. For example, in modern theatre and television it is becoming increasingly common to need to simulate computer interaction, particularly internet use. By creating a false replica of a well known website (such as a popular news site, or google.com), and altering a network so that traffic is routed to this website, it may be possible in a play to show a person accessing a well known website as if it was real - yet audiences will know that it is not, due to the way in which it is framed within a dramatic production.

In a trajectory this type of deception can be used on its own, purely as a piece of fun, and also as follows on naturally from a beat-the-system deception (see below for more on this combination).

Beat the System In this kind of deception, the computer is being tricked and this is used to trick the audience member who believes the computer's output.



Interestingly, this type of trickery actually happened with two of the participants in the Cup Game. However, it was reversed: in the section where they were trying to beat the computer, 2 of them actually switched the bead in such a way as to make the performer think that the computer was correct, meaning that the audience member had used the computer to fool the performer.

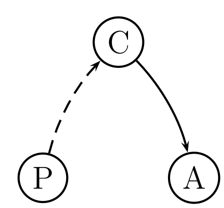
This kind of deception may be useful for performers who are using appropriated technology that they do not have access to. For example, by turning on the mirroring function on a web cam, it is possible to make oneself seem to disappear whilst using live chat software [137]. This is a classic magic use of mirrors for a disappearance, but updated by use of the web cam technology to become a remote trick rather than a stage act.

The difficulty with Beat the System deceptions is that sometimes a system will not have a useful weakness. They are also somewhat hard to perform live, in that often you will have to also fool the audience directly (as happened in the Cup Game), as they can see what you are doing. The flip side of these difficulties is that, as demonstrated by the Cup Game, people have a high level of trust in computer systems. If you can trick the computer as well as a person, this can increase the impact of the trick. Tricking a piece of software that is already known to the audience could add further to the impact. These deceptions also offer a lot of potential for remote trickery.

In terms of framing this type of deception is most likely to be used with the frame of a magic trick or story telling deception, to support a deception which is directly being performed on the audience member. This is because unless the illusion being created is purely that of something happening in the computer system, it will be used as part of a larger story that is being told.

In a trajectory this can be revealed to become a stupid-computer deception. It is also interesting to use this following on from a magic-trick deception. In this case, first the workings of the computer system enabling the magic-trick are revealed, then the weaknesses of this computer system are used to support a new deception.

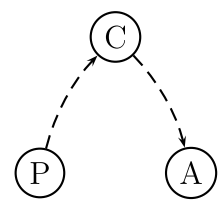
Caught by a Computer In this situation, the performer is trying to trick the computer, yet the computer knows this is happening, and tells the audience that it has happened. In the cup game, this is reversed, with an active participant trying to trick the computer, and the 'cheating alarm' letting the magician know that the participant is trying to deceive him.



In some situations this may be a useful part of a deception which can be applied to make a system seem more robust and unbeatable. Magicians often use technical devices that seem to stop trickery, such as fitting padlocks and chains around a box in which an object is to appear, and various ways are used to demonstrate to the audience how strong the chains are, and how they can't possibly trick their way into the box. This kind of deception is used to support a framing magic trick deception (which itself may be framed by a story telling deception). A second potential use for this in performance is to constrain the possible actions of a participant as in the Cup Game, by using the computer to catch deceptions attempted by the participant.

As part of a trajectory, this kind of deception is particularly useful before a beat-the-system deception, by showing how the computer is unbeatable before it is in fact beaten. This type of trajectory can also be reversed, a deception can be shown, then a caught-by-the-computer strategy being used to demonstrate how no trickery was possible.

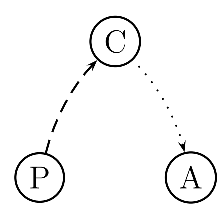
Computer Suspending Disbelief The juggling tracker is specifically designed to suspend disbelief when it receives an input which is out of normal range, and to interpret it as being something within the expected range of inputs. For example if multiple jugglers move in front of the camera, it will treat this as one juggler. Similarly, if more balls are juggled than it can track (currently 8), it will just ignore some of them and keep working.



In the juggling tracker, this suspension of disbelief makes for some interesting comic effects when multiple jugglers may appear to the system as one, creating the (obvious to the audience) illusion of a single very very good juggler working with the system.

One important thing that was noted with the juggling tracker, was that complex deceptions like this only really work once the audience are already clear about what the system is actually doing. This meant that a trajectory of effects including some of lower complexity were required before this would be funny for the audience.

Sneaky Computer Rock is also intended to interpret abnormal inputs as being some kind of meaningful inputs, for example if people pick up the rock, it will still detect movement, and respond in some way, even if the way it is responding is not very well defined at these points. This is designed so that it is not entirely clear to those interacting with the rock exactly what it does respond to and what it does not. This means that people can get a general idea of how the rock responds, but because they cannot tell which things it does not respond to, they are left unsure of the limitations of its sensing.



The Cup Game's In Front of Your Eyes trick also makes use of this kind of deception. The computer is told that it is about to be fooled. The cheating alarm is turned off, so it will not catch the performer cheating. The computer output is altered to remove camera video and only show a black background and red circles where the cups are, in order to avoid the trick being seen on the video feed. This makes it easy for the performer to trick the computer, and also provides a useful misdirection for the audience member. This also makes use of the expectations of the computer being honest that are set up by the first trick.

Another example of this kind of deception is a strategy used in the Juggling Tracker to handle drops in some performances. It can detect that it can't see any balls, but it doesn't know exactly why. Because of this, the script chooses to repeat the current visible output until it can see a pattern again. During this time the performer can kick the balls up in an impressive manner and act as if that is the trick. This is an exploitation of the audience's lack of expectations, as they do not know what trick is planned, and the system is designed not to notice failures.

For a performer, this strategy is potentially very useful. Having a computer which responds sensibly to attempts to trick it and out of range or unexpected inputs without being obvious to onlookers may be the difference between looking like a professional and looking like you are making mistakes all the time. This relates to the work of Benford et al. on making systems which handle unexpected inputs gracefully [10].

In installations like the Rock, making it respond interestingly to people trying to trick it or inputs that it does not expect can make it appear to have far more complex interactions than it really does and add to the complexity of people's interpretation of it.

This type of deception has been used in this thesis within the wider framing of other deceptions, as a way of supporting the deception and making the deception richer and more robust. It is hard to see how this could be used to frame other deceptions. In terms of trajectories, again like the computer suspending disbelief deception described above, it only really makes sense if people know already what the computer is supposed to do. For example in the Rock, the way it behaves when it receives an out of range input only makes any sense when it is integrated into people's existing interpretation that has been created by previous good inputs.

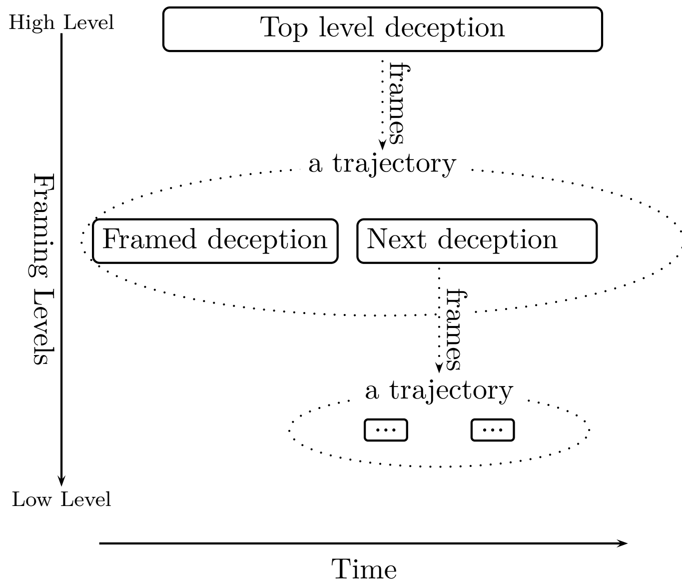


Figure 7.1: Multiple layers of deceptions

7.1.4 Trajectories and Framing: Multi-Layered Deceptions

This discussion has so far described trajectories and framing individually. However in real situations, both may be used. Trajectories and framing can be seen as different but closely related ways of combining deceptions within a performance. Essentially, trajectories relate to a temporal ordering of deceptions whereas framing describes a composition of one or more low level deceptions to support a higher level deception.

It is important to note that the a deception can potentially be part of a trajectory as well as being used to frame other deceptions. One way to view this is as a multi-layered structure, with framing deceptions at one level containing a set of lower level deceptions, and potential for trajectories between the deceptions at each level. Figure 7.1 shows the way a deception can frame a whole trajectory of deceptions, which can themselves frame lower level deceptions and trajectories.

This multi-layered structure of trajectories is considered here with reference to Benford et al.'s work on designing trajectories through experiences [9]. They describe several different types of transitions as key to the creation of a smooth trajectory through an experience - *critical moments in an experience at which* *users must cross between spaces, rub up against schedules, take on new roles, or engage with interfaces.* These are moments when the performance goes from one state to another, for example the point at which a person has been briefed at the start of an pervasive game and is about to become an active participant in the game.

With reference to the multi-layered structure described above, reveals become an interesting new example of a transition point, where one layer of deception is peeled away, meaning that the participant of an experience may now know how they were deceived. The person is not actually transitioning between a physical state or necessarily changing role, they are simply transitioning to a situation where they have greater knowledge, so are not being deceived any more (although they may still be subject to a higher level deception).

A classic example of a trajectory involving reveals is the magician's mistake or sucker trick, where a trick is performed in a way that makes it look obvious to the audience how it is done. However when the magician reveals the trick, it turns out to have been the setup for a much harder to explain trick.

The Cup Game is a very good example of a multi-layered performance using reveals. At first in the game, the person knows nothing about the computer system. This lack of knowledge is used to trick them. Then, a reveal is used, to show how the trick was performed, so that they now understand that level of the performance. However whilst the audience now have more knowledge, they do not know about the second trick which is hidden as a sub-action of the demonstration of the first trick. Later on, this trick is again revealed, by which point they are actually aware of the complete sequence of actions in the performance.

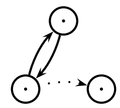
Reveals are also often seen as part of a trick itself. For example when a person thinks the bead is under one cup, the performer does a reveal, letting them know that it is under another cup. This alters the type of deception they are subject to, as it reveals that they have been tricked, whereas previously they were unaware.

Perhaps the inverse of a reveal is the creation of deceptions by setting up stories (a 'setup'). For example in the Cup Game, the first thing the performer does is tell a story that makes the audience member believe that the cup will be found by listening to the sound. This sets up a deception in the mind of the audience member, creating a transition from being undeceived to being subject to a deception.

(1)Introduction

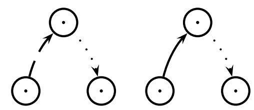
Setup the Cover Story

(2)The Cups by Sound



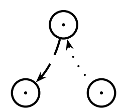
Reveal the existence of technology and how it tricked them in part 2

(3)In Front of Your Eyes



Reveal the bead is not under the cup they think

(4)Get Your Own Back



Reveal the trick played in part 3

(5)The Big Reveal

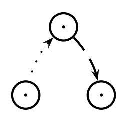


Figure 7.2: Reveals and Transitions in the Cup Game

Figure 7.2 shows in **bold** the reveals and setups that occur in the Cup Game. These are shown in time order as they fit into the trajectory taken through the performance. Figure 7.3 (overleaf) shows this same set of transitions in terms of how the cup game slowly increases the knowledge of the participant about the tricks and the computer system. In this figure, movements down the page denote an increase in knowledge - note that the first transition that occurs is actually the audience member being told a false story about what is happening - this setup is shown as a decrease in knowledge of the system.

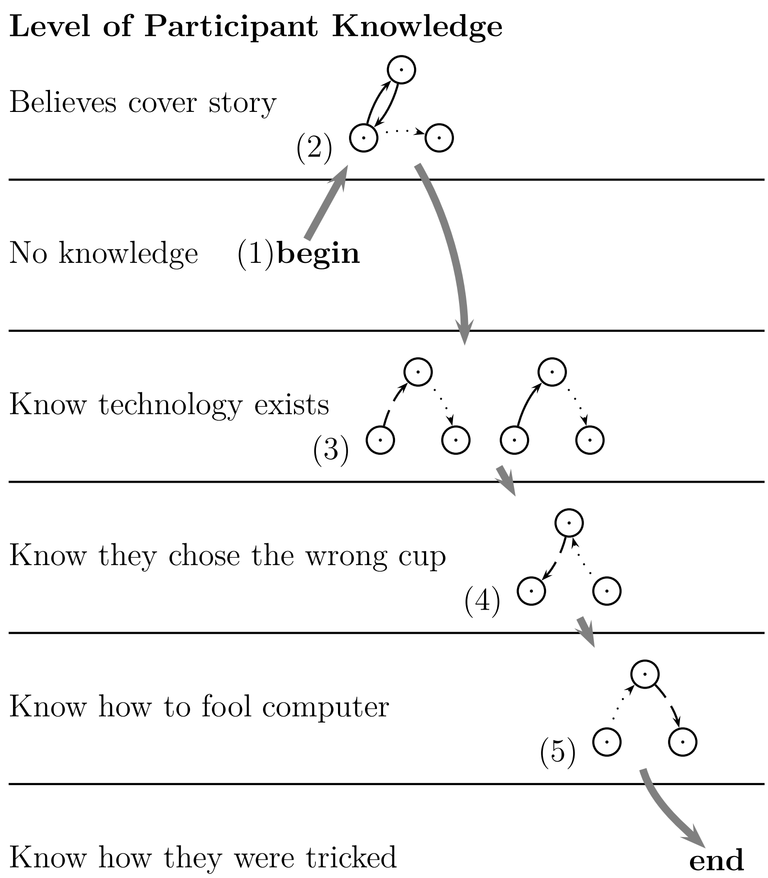


Figure 7.3: Layering of Reveals in the Cup Game

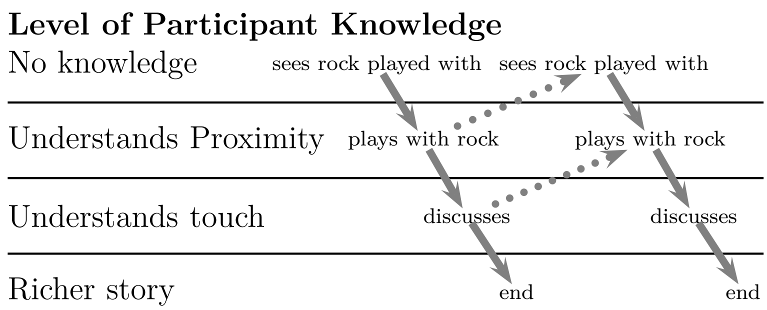


Figure 7.4: Interwoven Trajectories in Rock

Benford et al. [9] also mention interweaving of trajectories, where trajectories through an experience followed by one individual may bring them together with those of another individual into a shared trajectory (which may later diverge). These individuals may not be at the same place in their own personal trajectories - for example theme park rides design queuing systems so people queuing can see those who are already on the ride. Rock demonstrates an interesting use of these points at which trajectories interweave. Participants who are more experienced in interacting with the rock pass on their greater knowledge in order to encourage a transition for other participants. This transition effects participants in two ways. Firstly, it encourages touching of the rock as opposed to simply viewing it, a classic transition between roles as described by Benford et al. Secondly, it passes on the story about the rock, spreading the deception to people who do not currently know about it and allowing the creation of a richer story. Figure 7.4 shows two of these parallel trajectories, with links showing the ways in which the installation is made richer by the two trajectories coming together. We also need to consider points when we do not wish people to reveal things, for example in the cup game, it is desired that those who have previously seen the tricks do not reveal to later participants what is happening. When two participants knew each other, the first participant was asked to keep the secret; participants were happy to respect the secret of the tricks as they didn't want to spoil it for the others.

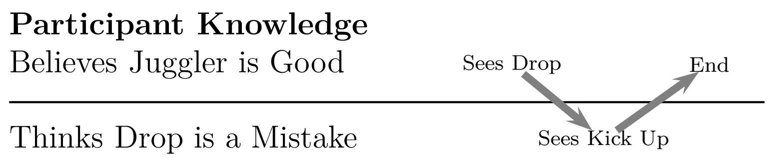


Figure 7.5: Telling Stories about a drop whilst juggling

A similar setup of a story occurs in the Juggling Tracker when a ball is dropped and a get out trick is used. This trick is intended to make people believe that the performer put the ball on the floor deliberately just so they could do an impressive kick up back into the juggling pattern. This creates a transition from a situation where the fact that the performer has made a mistake is accidentally revealed, to a story where the audience believe that he really is a highly skilled performer who can juggle as well with his feet as with his hands (stretching the truth somewhat). Figure 7.5 shows this transition. This trick often relies to some extent on a suspension of disbelief - the audience may know full well that it was a mistake, but they are willing to pretend otherwise due to the juggler showing that they are still in control of the situation.

Patterns

The different trajectories of each case study can be seen as patterns for the use of reveals and setups more generally. Four different general patterns are described here, based on trajectories observed in each of the three case studies. This is not a complete taxonomy of trajectory patterns, as clearly many combinations may be possible, however it does demonstrate the way in which the trajectory framework can be used to extract these more general patterns from a performance.

Reveal/Setup In the 'In Front of Your Eyes' Cup Game trick, the reveal of the first trick is itself used as a setup for the next deception. This kind of fluid transition from deception to deception is common in magic performance. This is an example of what Nardi [99] describes as changing the bracketing of a trick, where events that seem like they are something else are actually the beginning of a trick. This can be particularly effective when combined with the sense of surprise provided by the reveal.

Cover Up The juggling tracker uses a classic strategy to cover up mistakes by performing a virtuosic action beginning from the error state which gets the performer back into the core of the performance. Similar strategies are often used by live musicians to cover up wrong chords or notes by making them part of a progression that ends in the correct place, making it seem like a flourish rather than a mistake. This can be seen as first an accidental reveal of the fact that the juggler is not perfect, followed by a setup of a story, essentially a reveal/setup but performed in the opposite order.

Onion Skin This strategy, based on the cup game, uses successive reveals in order to let the audience know more and more about a performance situation. At each level, an illusion is generated, based on what the audience does not yet know, which is then revealed by the next level. In this way, the audience can get deeper and deeper knowledge about a system, whilst keeping an interesting level of illusion at all points. The reveal/setup pattern described above can make this strategy more fluid than splitting things explicitly into individual tricks.

Telling Tales The strategy used in the Rock uses people who are only slightly more experienced in an interaction to set up both the story of the performance, and also to help reveal practical aspects of the interaction. This is in contrast to the typical method of creating an interactive system and using pre-taught experts in order to aid those wishing to experience it. This has much in common with the communities of internet videos and comments that have sprung up amongst musical learners, which have enabled many people to learn musical instruments (or other hobbies / sports etc.) without expert tuition. It also has much in common with the word of mouth traditions of folk storytelling and music, where stories and songs evolve based on the input of those passing the story on, allowing for a richer, more diverse range of performances than those envisaged by the original writer. This particular pattern may also be something to be actively avoided, as with the cover story in the cup game, where it was important to avoid participants passing on the revelations made during the performance. In some situations, it may be that we only wish people to pass on a setup, but not the reveal, so people know what the story is, but the surprise of the reveal is not spoilt (this is a bit like Reeves et al.'s 'Suspenseful' strategy [109].)

7.1.5 The Deception Framework Evolved

In summary, the deception framework involves three things. Firstly, a concise description of the perceptions and understandings of the people and computer systems in a performance. This is inspired by the work on public interaction and spectator interfaces of Reeves et al. [109]. It allows for a succinct description of what people are able to perceive in a performance, and what is not visible to them.

Secondly, deception can be used to create illusions, or situations in which what people believe they know about a situation differs from the truth. A notation and several different strategies for creating these deceptions are described. This also takes account of the fact that many ‘deceptions’ in performance actually involve a suspension of disbelief, where those being deceived are fully aware of that fact. This work builds on prior HCI work on ambiguity [50], interpretation [116] etc. by applying insights from the literature describing professional magic performance.

Finally, as described in this section, when practically applying the framework it is important to consider is the things that come between and around a deception. This includes the way that deceptions frame and are framed by other deceptions and also the way in which deceptions are put together in trajectories. In particular, Benford et al.'s trajectory framework [9] is extended with two new transitions, a setup, in which a deception is created, and a reveal, in which the deception is exposed.

The next section will explore further ways in which the framework may evolve in the future, by using a more broad definition of performance which allows it to address a wider range of situations.

7.2 Extending the Definition of Performance

This section explores potential areas for future development based on extensions of the definition of performance used in this thesis.

Firstly, the case studies primarily involved roles of performer, audience, and computer system. It is possible to consider more complex performances, with multiple different roles with respect to the framework.

Secondly, the extension of the framework to performances in which the performer is separated from the audience by space or time is considered, as opposed to the synchronous, co-located performances studied in the core of the thesis.

Finally, the use of these techniques in non-performance settings and applications is explored. The last 4 parts of this section explore key HCI areas where performance-like interactions may occur. These performance-like situations have enough in common with artistic performance for these deceptive techniques and the creation of illusion to be equally useful, although potentially more care is required when using deception due to the often unclear framing of situations.

7.2.1 Performances with Many Roles

Rock showed how some performances may have changing roles, where people change role through a performance, and showed a complex pattern of deceptions, where a deception was passed on from one person to another. In some performances, there may be multiple roles or types of people involved, all of which may be involved differently in deception. For example, Stelarc's Ping Body performance [160] (described by Dix et. al [39]), involves two audiences, an online and a co-located audience, who each have different knowledge about the performance and are able to view different aspects of it. Similarly, multiple different categories of performers may by involved in a performance in addition to the obvious performers, such as behind the scenes orchestrators [78], lighting operators or other back-stage roles. It is relatively easy to extend the deception framework to support these performances simply by considering all the possible deceptions in any particular situation. In these more complex performances, the many overlapping perceptions may be significantly more complex to analyse, and the number of possible deceptions may be significantly greater. Whilst this makes things harder in some ways, it also creates a greater potential for deceptive techniques to be useful.

7.2.2 Remote and Asynchronous Deception

As demonstrated by magic on television, youtube.com or other similar video sharing sites, it is possible to translate stage and live performances of deception into a remote performance with some success.

However, there are some ways in which the use of communication technology to mediate a performance may significantly alter deceptive possibilities. When the performer and audience are remote, the default is no longer that the performer is visible to the audience. Instead, only what the performer chooses to make visible to the audience member is visible to them. There will be constraints on how much the person can hide however, due to the medium of communication used. For example if a person uses video conferencing as their mode of communication they will be expected to make their face visible on the camera (this expectation, and the need to make one's face presentable, is often cited as a reason for the repeated failure of video phone technology [122]). As long as they work within the constraints of the technology, performers may gain a great flexibility to be able to deceive. For example in TV shows, backgrounds are often added in at a pre-processing stage, meaning performers can seem to be in a particular setting, whilst actually they are in a studio.

As well as altering the ability to deceive, the use of a remote communication method may also alter the ability to detect audience responses to the deception and to respond to their actions. For example if performing magic on a web cam it may be impossible to watch people's eye movements in order to tell whether they are looking in the right place.

In terms of audience responses, asynchronous media, in which viewers seeing data at some point after it has been recorded, and broadcast media, in which many viewers see the same data may also create different challenges for performers. With a broadcast medium like television there is no real-time response; even in live TV the performer has to assume that the audience has been tricked, and cannot respond to them in any way. In most TV magic, this is avoided by making the viewer a passive participant and having active participants in the studio or on location wherever the magician is. Other media, like email, postal communication and web site postings are essentially asynchronous, with viewers seeing data at some point after it has been recorded. In these media performers may have access to responses after some time but similar issues exist to those of live television.

Asynchronous media also create opportunities to change the temporal order of things. For example on TV programs where they trick people, people may be asked for permission to show the footage to an audience at a point after they have been tricked. Thus, the actions are performed to an audience after permission has been given, yet are performed on the victim before they know about the trick.

Two planned pieces of future work aim to explore remote deceptions in synchronous and asynchronous settings. In particular it is planned to explore the performance of magic first via web cameras, mobile phones and internet chat (in real time), and secondly via email and the web (asynchronously), with an aim of designing tricks that exploit the nature of these technologies in order to create magic that could not be done in person.

7.2.3 Social Deception

So far, this thesis has considered deception in co-located synchronous performance settings. Clearly deception does occur in non-performance settings, and may even be desirable. One particular area where deception does occur is in social settings. Studies of social interaction suggest that up to a 3rd of social interactions involve some kind of deception [35]. This is not necessarily for negative reasons and may be for reasons such as making people feel good, e.g.*'Told her her muffins were the best ever'* [35]. Designers can design with this kind of deception in mind. For example, Hancock et al. [56] study the use of various technological communication methods and demonstrate a significant effect of the design of a medium on how and how often people lie. Aoki and Woodruff [4] describe ways in which mobile phone or instant messaging users may use the lack of information in the communication channel to support stories about a lack of response to a communication attempt, in order to avoid the other user feeling hurt or feeling a loss of face. Interestingly, they even talk of what is effectively a suspension of disbelief on the part of the person being tricked:

> *Whether or not the stories seem "transparent" is not (necessarily) the point. Such explanations are often transparent in that sense; preservation of face requires credibility and the willingness to accept an explanation, not actual belief [4]*

In this situation, something very similar to a performance deception is occurring; someone is performing the role of being unresponsive due to some technical limi-

Figure 7.6: A lag switch for cheating in games

tation (such as claiming to have left their phone at home), when in fact they are unresponsive by choice, but do not say that in order to preserve social harmony. Boehner & Hancock [16] even suggest various rather extreme ways in which designers might support this kind of performance, for example by making 'phones with a personality', that arbitrarily decide to hang up calls or not to ring. This is essentially designing stronger expectations of a failure event into the interface, in order to support a particular type of deceptive performance.

These designs have explored story telling and magic trick deceptions, where the user is controlling the system in order to perform the deception, which is taken at face value by the target of the deception, potentially with some suspension of disbelief. There is a spectrum in this work from designs which do not support deception explicitly, such as email systems, to Boehner and Hancock's designs to aid deception, through to systems such as the Swarm system described in Chapter 3 [114], which makes deception an explicit part of the interface.

It is also possible to imagine the use of trickery of the computer itself to achieve similar effects of social deception. Two examples of this would be sending false data to a piece of location aware software, in order to make it seem like you are in a different place, or by deliberately degrading network communications in order to make your deceptions harder to detect (as higher bandwidth communications may make deception easier to spot [63].) In computer gaming, this kind of deception has already been used (for negative ends). With some games buffering up player actions during a network delay and then sending them all as soon as the network comes back, the use of a lag switch which attaches to the network connection to cause brief delays can give players a big advantage (Figure 7.6).

As social networking and websites primarily based on user generated content become increasingly part of our normal lives, the possibility of performing ourselves in these mediums will become particularly relevant. Designers of such systems may find it useful to consider how these tools support illusion.

As well as people telling white lies with the help of or via computers, creating lying computers may be useful too. Cristiano Castelfranchi suggests several reasons why a computer lying to a person may be positive [26]. For example if a computer system or its designer knows that a person doing a presentation is likely to run over time, it may give them a warning when they have 5 minutes left; he argues that for people who are persistently late even with this cueing, the computer system may be programmed to give them the 5 minute warning with 7 minutes left. He also argues that in some situations, computers will be used to support the self interest of the person controlling or designing these particular computers, which may make it desirable for them to deceive other people, for example to deceive a malicious user in order to keep a system secure.

One interesting way in which a system can fool a person is when they do not even know it exists. For example in a close circuit TV surveillance system at a major UK railway station, random movements of the cameras have been implemented, even while no-one is watching that particular camera. These make it seem to people watching the cameras as if a person watching a camera is actively tracking passers by, as in a crowded situation, the random movements will inevitably match at least one ongoing movement in the space. The typical viewer will not even realise that a computer is behind the camera's movement. A more technical viewer may infer from the camera movement that some kind of automated tracking system is involved. In both cases this technology is deceiving the viewer into thinking there is something or someone actively viewing the footage.

7.2.4 Non-Entertainment Performance

Whilst social interaction has many performance-like qualities, many more formal interactions are performances, even if they are not performances for entertainment. Corporate presentations are a classic example of this and many performers have made use of this fact to sell their services as presentation skills coaches. For example, magician James Freedman teaches presentations skills and sales tactics using the psychology of magic performance in order to teach people to direct attention, perform misdirection and other things useful to those presenting a particular message. Other non-entertainment performances include those of politicians, lawyers, advertisers and teachers. Even announcement systems in shops and on public transport have performance elements. It has been argued that successful performers in these contexts use techniques similar to those of magicians, even if they use them purely in ways sanctioned by society [99].

One very simple example of the benevolent use of deception in a non-entertainment setting is the mysterious Inspector Sands, who is often paged on the public address system at London railway stations. This is in fact a coded call that a fire alarm has been set off, in order to avoid causing general panic whilst it is investigated. Similar deceptive codes are used in shops, theatres, etc. where a message must be sent only to certain recipients by a person who is performing to all those in the building. Reportedly, in the London Underground case the messages are actually produced directly by the fire alarm system itself, meaning that what is essentially an automated alert is sent out under the cover of what a normal person will think is a live public address system [102].

7.2.5 Public and Ubiquitous Interaction

HCI authors have discussed the idea of everyday interaction with devices or systems that occur in public being treated as a performance [109] and of using the performance of everyday living as a design tool [67]. Essentially, this considers any witnessed interaction, where someone else sees the interaction or the results of the interaction, as being performed to the witnesses. Interactions and their results are increasingly occurring in front of audiences in multiple ways. Firstly, fixed public interfaces are becoming common in museums, shops, galleries, on public transport etc. Secondly, mobile computing devices such as smart phones, mp3 players and laptops are becoming increasingly common and the range of uses is expanding fast. These devices are often used in public in front of others. This study of public interaction as performance also makes sense due to the increasing number of performance and artistic projects that use publicly sited interfaces or mobile devices - many challenges may be similar for both artistic and non-artistic projects.

Interactions in public settings create many interesting challenges. These interactions are performed in front of others who may be knowledgeable audience

Figure 7.7: Joe Malia's private text message scarf [152]

members or simply unwitting bystanders. There are many ways in which performance may be a useful model for these situations, for example when we wish to encourage bystanders to become active participants [118, 126], or to consider the effect of not knowing whether bystanders are actually a knowing part of the performance [8], or ways in which we may support personal privacy whilst siting interfaces in public or semi-public places (e.g. [14, 76, 24]). Figure 7.7 shows a humorous example of a special scarf designed to allow people to send text messages from mobile phones in privacy whilst in a public place.

An interesting example of the use of deception in this kind of performance is the use of mobile phones in public. It has been shown that bystanders to a mobile phone conversation will adapt their behaviour, for example by turning away from the caller in order to seem as if they are not listening to the phone call. However, the same study demonstrated that these 'non-listening' bystanders actually had very good recall of the content of the conversation [85].

7.2.6 Computer Security

Whilst the intentions of the work described in this thesis are very different to those of computer security experts, several of the ideas and strategies presented in this thesis are in part similar to those used in the field of computer security. This section briefly describes some of the work on deception in computer security, and how performance-like elements can be found in security situations.

Deception is used both in defending and attacking computer systems. For example, in one high profile case described by Stoll [124] a person breaking into a computer network did this by pretending to be a legitimate user, whilst people defending the same computer system deliberately allowed the intruder to think that no one had noticed the intrusion, in order to see what the person breaking in was actually doing. This is essentially a suspension of disbelief by the people defending the computer system, controlling the computer to perform a role of an unprotected system, whilst taking care to avoid letting the intruder cause real damage. Computer security experts also describe strategies where a person is tricked in order to gain access to a computer. These are known as social engineering [95], and are used because it is often easier to trick a person than to directly break into the computer. These strategies again rely on a significant performance element. For example people may phone workers posing as being from the company IT helpline in order to get passwords from a victim.

One interesting recent development in computer security is the use of inconsistency in designing defensive deceptions [100], which is in many ways analogous to the use of ambiguity in design, removing the fixed interpretation of what is occurring that would be available with a single response. There is also some work on designing computer systems which automatically modify their deceptive responses in order to pick the one which will most confuse an intruder [112].

There is evidently a cross-over between the use of deception in performance and the use of deception in studies of computer security which has scope for future work. A performance interpretation of computer security based on the deception framework outlined in this thesis has potential both to usefully inform the threat focused work of security researchers, and to allow insights from computer security to aid those using deception in performance.

7.3 Ethics of Deception

In some performance situations there is clearly no problem with deceiving the audience, as this is expected or implicitly permitted by the overall framing of the situation. In some settings however, deception may need to be considered more carefully. This section considers the ethical issues related to the use of deception, and potential approaches to resolving these issues.

7.3.1 Potential Harm of Deception

There are several potential ways in which deception may be harmful to others. These harms can be related to the deception itself, and also the reveal, examples of 4 possible types of harm are described here.

Effects of deceiving A deception may cause harm to participants directly. For example, a false medium may use deceptive techniques to cheat people out of large quantities of money.

Effects of non-revelation Not revealing a deception may also cause harm. In the false mystic example, the continued non-revelation of the deception can increase the potential for harm greatly, as it leaves the person likely to believe other false mystics.

Effects of revelation Revelation of a deception may also cause harm. Again, in the medium example, the discovery that someone has not in fact been talking to a dead relative and has been being fooled may cause psychological harm to the victim.

Effects of not deceiving It is possible to argue that there are some situations where deceiving someone is the least harmful option. Some of the social deceptions discussed in Section 7.2.3 may be caused by people not wishing to harm others. For example, in the case that one is asked the question "*Do I look fat in this?*" many would argue that it is okay to say no. However, even this common kind of deception may be harmful in the case of a revelation; in this case, if someone is later told that they do in fact look fat, they will feel deceived and hurt, as someone they trusted was not able to tell them the truth.

7.3.2 Do People Know they Are Being Deceived?

Knowledge of the deception, or whether deceptions are likely is clearly an important factor. This section describes 3 possible situations that may be the case when a deception is being used.

People know they are being deceived In some situations, such as theatre, people actively know that they are being deceived, yet suspend disbelief in this.

People know they are in a performance setting (or other setting where deception is likely) Again, if people know they are in an established performance frame, such as at a show or watching a street performer, they know that deception is possible.

There is no knowledge of deception If there is no performance frame surrounding the deception, people may have no idea or expectation of deception.

7.3.3 Approaches to the Ethics of Deception

This section discusses two professions which make extensive use of deception in their work; magicians, and experimental psychologists. They take very different ethical approaches to deception, which are well documented. These contrasting approaches and the justifications given for them are described here.

Magicians Since at least the 1890s, performing magicians have had a relatively straightforward approach to deception which is that deception is ethical only if it is framed within a performance situation, where it is made clear that the performance is a trick. This relies on what Tognazzini describes as the audience's "tacit understanding" of the deception:

"*A magician who forces a seven of diamonds within the confines of a magic act is acting ethically: He and his audience have a tacit understanding that he will be carrying out sleight-of-hand. The same magician who later that night forces the same card while playing a game of poker is a cheat.*"[128]

Magicians such as Davey in the 1880s [33], Houdini [65], and more recently The Amazing Randi [107], have worked to expose those who use the techniques of stage magicians but claim that they have paranormal powers. They see this use of deception as exploitative and immoral. This has not always been the case - according to Triplett [129], this emerged in tandem with the growing professionalisation of stage magic acts in the 19th Century. Prior to that, the boundary between magical entertainment and supposed real magic was much less well defined.

Psychologists In philosophical terms, the magician's ethics of deception are essentially a deontological position - stating that the use of deception when framed as performance is acceptable, but that the use of unframed deception without people's consent is wrong, no matter what the outcome. In contrast, psychologists typically take a more consequentialist view of deception, seeing it as being acceptable or not based on what the consequences of the deception are. This has led to many psychologists using deception in studies. Most controversially the famous obedience experiments of Stanley Milgram, in which subjects were deceived into thinking that they were giving fatal levels of electric shocks to a person (whereas in fact the person being shocked was revealed to be an actor faking the pain [94].)

This and other controversial experiments led to a debate amongst psychologists as to the use of deception, with writers such as Kelman [74], arguing that restrictions on the use of deception were desirable. At present, psychology generally does not see deception as a positive thing, but as a justified evil - as demonstrated by the British Psychological Society code of conduct:

> *Avoid intentional deception of clients unless: (a) deception is necessary in exceptional circumstances to preserve the integrity of research or the efficacy of professional services; (b) any additional safeguards required for the preservation of client welfare are specifically considered; and (c) the nature of the deception is disclosed to clients at the earliest feasible opportunity. [19]*

The American Psychological Association has a similar description in it's code of conduct [1], which was first drawn up in direct response to the Milgram scandal. These codes of conduct define the use of deception as a last resort. In practice, it is clear that the range of situations in which it is deemed to be necessary is wide, in one study, 43% of the social psychology articles surveyed described studies using deception [60]. In fact, the existence of several very similar studies in which some use deception and others do not, suggests that the use of deception for many experimenters is more a convenience than a necessity [61].

The primary argument in favour of the use of deception in psychological studies is that the good of doing research may outweigh the risks of the deception:

> *[the] potential loss of important research benefits from the decision not to do a study needs to be weighed as seriously as the risks involved in doing the study [77]*

There are also secondary arguments, as to whether participants are actually harmed by deception, with some even arguing that participants enjoy deception studies more [27].

A distinction is also drawn by psychologists between deceptions which are clearly part of an experiment, and deceptions which go outside the framing of the experiment. For example some psychologists have studied the actions of people who believe that they are performing experiments on others, or people who think they are in a waiting room waiting for an experiment to begin. Deceptions within the experiment are often seen as less serious or potentially harmful than deceptions as to the fact or timing of the experiment itself [74].

7.3.4 What Does this Mean for Performance?

Even using the stronger magicians definition of the ethics of deception, most performance uses of deception are perfectly acceptable. Key to this is the framing as a performance, and the framing of actions within it. In the situation that people are aware that they are taking part in a clearly framed performance, hidden deceptions within that performance are very likely to be acceptable.

At the point where we break away from well understood performance frames ethical issues become significantly more complicated. For example, in pervasive gaming, onlookers may seem to players to be part of the game, yet they do not know that the performance is occurring at all [78]. If a theatrical situation is being played out in public, care has to be taken if there is a risk of onlookers being also deceived. A good example of this risk is Blast Theory's Kidnap [141], where a person volunteered to be kidnapped at an unknown time and place - this required careful planning in order to avoid the police being called or bystanders being distressed seeing someone being kidnapped [105]. As well as the risk to onlookers, deceiving participants where the performance frame of a pervasive game or performance is not clear may cause problems - for example if performers are not known to the audience member it is unclear whether they are part of the performance, which may make their actions worrying to those taking part in the performance.

Even within established conventional performance frames such as theatres, deceptions which break outside the performance frame may be ethically problematic. As performers play with and push the boundaries of traditional notions of performance staging and framing, it may become increasingly unclear whether events are occurring within or outside a performance.

In situations where the magical definition of ethics clearly does not support the deception, it is hard to see a justification for the use of deception in performance. Given the primary entertainment purpose of performance, artists and performers do not have the psychologists' justification of the potential great worth to humanity of their work to allow them to deceive in these situations.

Application of this kind of approach to deception may also be professionally useful for performers. Stage magicians developed their code of conduct throughout the 19th Century as a way of separating themselves from charlatans, con men, false mediums and others who exploited similar techniques. With the growing popularity in interactive art of new technologies such as surveillance cameras, tracking technologies, ubiquitous computing, all of which may generate suspicion and mistrust in the public, the taking of a strong ethical stance on deception may help interactive artists and performers to avoid being tarnished by those who use the technology for more negative purposes.

7.4 Conclusion

This chapter discussed the deception framework in detail, with reference to the performance experiments described in the previous chapters. Two concepts are identified as key to the framework, that of framing a deception within a wider performance setting (which itself is a type of deception), and the idea of trajectories of changing deceptions (and potentially revelations of these deceptions) over time.

The potential for extending the definition of performance in the framework is also discussed. The framework can be extended to consider more complex performances, with multiple roles, or with parts of the performance occurring in different places or at different times. It can also be extended to more general HCI situations which whilst they are not artistic performances exhibit performance-like aspects.

The final section describes ways in which ethics may constrain the use of deception, describing the contrasting ethical approaches of magicians and experimental psychologists, both professionals who often use deception in their work. The more constrained ethical position of magicians, that deception should not be used outside the framing of a performance, is described here as being more suitable for use by performers than the approach taken by psychologists.

Chapter 8

Conclusions

This chapter considers the work in this thesis with respect to the stated aims of the work described in the introduction. In particular, the introduction suggested several questions that the thesis would answer and contributions which it would make. This chapter discusses the way in which the thesis addresses these. This is followed by a section discussing the potential future implications of this work for HCI, performance and wider computer science.

8.1 Back to the Beginning

In the introduction, the thesis claimed to address a set of questions about performance interaction and to make four particular contributions. This section revisits these questions and contributions with respect to the work in the thesis.

8.1.1 Questions about Performance

What do we wish people to see and understand about a performance?

It is clear from the literature described in Chapter 2 and 3 that we do not necessarily desire audience members in particular to understand or see absolutely everything that is going on in a performance. This is even more evident in the case studies; even in the juggling tracker, the most straightforward of the systems developed here, several situations emerged where it was desirable to hide the communications between the performer and system from the audience. In the other two studies, things being hidden or not understood by the audience were key to the way in which they worked.

It is not just desirable to hide things from the audience. For example in installations and other systems where people effectively become performers by interacting with a system, obscuring elements of the way it is working may create a richer experience. Even in situations with an expert performer it may be desirable to hide things from them, presenting more of a conversation with a system than a purely controlled system [157]. In the Cup Game, people were even encouraged to try and trick the computer, creating illusions in the computer's model of the situation. Hiding things from computers may become increasingly important as computers become better able to sense and understand the outside world.

When might we wish to use deception in performance?

As we saw in the juggling tracker, in performance situations where the performer is typically in full view of the audience, there may be a disconnect between what we *wish* the audience to see, and what the audience *can actually* see. It may not be possible to remedy this by physically hiding actions without compromising the rest of the performance, or it may be that what we wish the audience to see is actually physically or technically impossible. In these cases, the deceptive techniques described in this thesis may be useful in order to create the illusion that the audience are seeing what we wish them to see.

The notion of suspension of disbelief is also important. As demonstrated by the Rock case study, suspension of disbelief can be a powerful aid to performers by providing many of the advantages of normal deception, but without requiring the complete creation of illusion that is required to really deceive people.

What strategies might be used to achieve this deception and what social and ethical constraints are there on them?

As demonstrated by the case studies, the diversity of potential performance situations calls for a diverse set of deceptions. This thesis describes several different strategies for creating deception - ranging from creating suspension of disbelief by telling stories to using technology to perform magic trick like deceptions, or even tricking the technology itself to support an illusion. Each of these may be suitable in different situations.

Many performance situations may require elements of multiple types of deception, and sometimes even deceptions contained within other deceptions. Designing the trajectories through deceptions during a performance is key to this. In particular, managing when and how things are revealed to people is highly important.

In terms of when it is acceptable for deception to be used we can learn from magicians, who have used deception in performances for many years. They have a clear and simple rule for use of deception which is that it is okay to use deception in a situation where you are clear that it is within a performance. This tried and tested ethical stance is applicable to most potential performance uses of deception, allowing for a wide range of deceptive practices, whilst avoiding misleading participants in a way which may be distressing, disappointing or harmful. By using this definition, it is possible to make use of the lessons of an established performing practice that has developed over hundreds (if not thousands) of years.

8.1.2 Evaluating the Contribution

The thesis defines a new perspective on interaction for the HCI community; that of considering the use of deception and the creation of illusion in interactive systems. This extends existing notions such as ambiguous design, design for multiple interpretations and design for spectator interfaces and framing.

The deception framework extends prior work by introducing the notion of deception as a useful interaction technique. Deception is used to alter the effective perceptions of a person in the performance, making them believe in an illusion, a situation which differs from reality. The illusion they believe in may not be possible to create in reality or may be impractical to create. Illusion has a further advantage over reality in that it can be supported by suspension of disbelief on the part of the person being tricked, making use of audiences' imagination and desire to be entertained to tell a story effectively without having to create wholly believable illusions.

In relation to prior work, the concept of perceptions to describe what people see and understand in performances extends work on public and spectator interfaces (e.g. [119, 109]). Deception diagrams extend the framework used by Dix [37] to represent collaboration, adding the possibility of describing deceptive communications and suspension of disbelief in a computer augmented performance. These deceptions build on notions of designing ambiguity [50] and multiple interpretations [116]. Finally, this thesis extends the trajectory framework of Benford et al. [9] with new transitions based on creating and revealing deceptions.

The thesis provides a framework and strategies for designing performances that involve deceptions. This is designed to be practical for use by interaction designers when creating real applications. As performances typically contain more elements than just deception and hiding, it is envisaged the framework may be used in tandem with previous HCI frameworks, design guidelines etc.

The idea of perceptions and actions as described in Chapter 2 provides a broad description of aspects of performance covered by several existing frameworks. This is used to allow the deception framework to build upon these frameworks, and potentially to be used in tandem with them. For example, the basic hiding of the interface in the Cup Game was considered with the public interaction framework of Reeves et al. [109] in mind, and Rock makes use of the idea of ambiguous design [50], and the suppression of obvious interpretations [116].

The thesis provides case studies of real-world designs using these concepts. These are intended to be useful to interaction designers wishing to apply the framework, and to artists wishing to build or use similar systems.

In this thesis, three case studies have been described which make use of the techniques described in the deception framework. These are all practical real world performance systems and have all been performed with in front of public audiences - at least 800 people have seen performances using the systems developed during the creation of this thesis. The systems described here were deliberately chosen to be extremely diverse, from the staged performance of juggling, to close up magic, and finally the gallery setting of an artistic installation. This wide range of case studies demonstrates the generality of the deception framework and its potential to be applied to a wide range of performance settings (and even potentially to a range of performance-like settings as described in Chapter 7). Two of the case studies (Rock, and the Juggling Tracker), have already been published in the HCI and performance literature.

The novel computer vision algorithms developed during the process of this work are of general applicability; it is hoped that these will be useful to those developing vision based applications in the future.

The pressure sensing algorithm used in Rock has also been published as a technical paper. This has led to some interest in using the algorithm from non-performance users, such as those developing medical technology. The algorithm is currently being developed further as an undergraduate project, aiming to apply the method to more general computer interaction. This will require the development of specific pattern and touch gesture recognition. This works on top of the system used for the Rock, which only required a continuous value for finger pressure, and did not perform additional analysis.

The tracker algorithm used in the juggling tracker is of general interest for tracking any of the many situations where people are interacting with objects. It was adapted and simplified for use in the cup game. A generalised version of this algorithm was also published in the British Machine Vision Conference.

8.2 Deception: Future Implications

The work done in this thesis has potential to be useful and have interesting implications for a wide range of settings, which are discussed in this section. The discussion begins with the areas of performance and computer vision which have been the core focus of the work here, and then discusses the implications for HCI more generally and for the wider computer science community.

8.2.1 Augmented Performance

A key implication of this thesis is that adding interactive technology such as computers into a performance situation is not simple, particularly when the system has it's own visual outputs. The added complexity is two-fold. Firstly, extra outputs created by the computer system mean that the audience's attention has to be divided between the performer themselves and the computer outputs they are creating. Secondly, the task of the performer is harder, as they have to both control the computer system and also interact with the audience. The deception framework offers a set of tools to manipulate and direct the audience's attention which can be useful to make audiences pay attention to the desired part of a performance or even to convince them that something has happened that is actually just an illusion. Manipulation of attention may not even need to be hidden, as in many performance situations people are willing to suspend their disbelief and go along with a deception even when they know that they are being tricked. This allows the performer to tell stories in a powerful and flexible manner, without having to create a perfect illusion to represent the story they are telling.

8.2.2 Computer Vision

The way in which computer vision algorithms were designed for these performances differs significantly from a traditional computer vision approach. Rather than designing an algorithm to be as good as possible based on a numerical error calculation, the algorithms here have been designed to be good enough to do the task, but with consideration to what they do when they fail. In some ways, the algorithms were actually required to be ‘bad enough’, with a level of failure that allows them to be tricked to create interesting effects, for example in the Cup Game, at some points the algorithm was deliberately designed not to detect certain performer actions. Graceful failure of algorithms is an interesting challenge which overlaps performance and algorithm design - for example in Rock, the pressure algorithm was designed to fallback gracefully to just hand tracking in the case of a pressure sensing failure. Even a level of unpredictability may be useful, as shown in the rock, where algorithm failures actually added to the implied character and aided the story that was being created amongst those interacting with the rock. Essentially, this work suggests a method of algorithm development where ‘goodness’ of an algorithm is not a simple numerical quality which can be easily calculated, rather it is dependent on the application in which the algorithm will actually be used. Several factors must be taken into consideration, such as how reliable the algorithm needs to be, how accurate it must be, how easy it is to trick, how it may defend against deliberate trickery, what it does in the event of a failure etc. This is less of a simple optimisation of a single variable, and more a balancing of algorithmic properties to create the right algorithm for the situation. Analysis of situations using the deception framework and the notion of computer and human perceptions provides a useful way to explore how to design an algorithm taking into account the multiple people who will be involved in a performance, and ways in which the algorithm design may support illusions being created in a performance.

8.2.3 General HCI

As discussed in the previous chapter, performance and performance like situations are common in modern HCI, particularly in the areas of mobile and ubiquitous computing, where computers take a part in people's lives that is interwoven with their day to day life. In these situations the deception, telling of stories and suspensions of disbelief that have been demonstrated in general social interaction are made more complex by the addition of technology. The deception framework offers a way to design, categorise and describe the situations which occur. As well as ubiquitous computing, social interactions are increasingly mediated by social networking systems, email and other electronic communication tools, creating more situations where people 'perform' their actions to others via computer mediation. Finally, a growing number of interactions are taking place through the mediation of static layers of technology embedded in our environment, such as video surveillance systems, public address systems, movement sensors for lighting or security purposes, which again are embedded in people's social situation. The work in this thesis will be useful in these situations. The deception framework allows the study of deception when computer systems are situated in performances and other social situations. The design of systems which take this deception into account will become increasingly relevant as complex computer systems continue to become a part of our daily lives.

8.2.4 Wider Computer Science

The field of computer security makes extensive use of deception, suspension of disbelief and ambiguity, which offer potential to be studied with the deception framework in this thesis. This link may also allow work from computer security to help inform HCI and performance work that involves deception.

As well as the potential for using the deception framework to study computer security situations, other areas of computer science involve a level of deception or suspension of disbelief. For example, in software engineering, software is often split into components which are designed to be treated as 'black boxes' which perform a particular task. Other people can call the component to perform that task, without worrying about what the component does, or how it does it. It may even be useful to hide the way in which the component works for a user, in order that they do not rely on it using a particular method. The original component can be altered as long as it still performs as specified. However in practice, component based design often breaks due to interdependencies between components at different levels. This can be conceptualised as the component developer creating a component which gives the illusion of performing a particular task. When the component based design breaks, the illusion cannot be sustained and the real workings of the component become a problem. Aspect oriented programming [75] attempts to fix this, by exposing 'cross-cutting concerns', aspects of a system which are relevant at many levels, for example security checks in financial software. In this case, what was previously hidden behaviour within several components, is taken out and exposed as a separate concern, so that whilst the external user may still not know fully how a component is working, they can see how relevant aspects of the component are handled. This allows developers to maintain as much of the suspension of disbelief about what the component does, whilst revealing vital areas of the component's workings. The application of the deception framework to software design problems could allow software engineers to explore exactly what they reveal and how they reveal it - aiming to achieve a maximum level of separation and re-usability of code (which is more likely to be possible if the user does not rely on how a component works), whilst still passing on enough knowledge to be able to address cross-cutting concerns and the dependencies they create.

8.3 Conclusion

This thesis has presented the idea of using deception in order to create illusion in the minds of those being deceived, and demonstrated the usefulness of these techniques within the design of computer systems for performance, and the design of performances using these systems. As demonstrated here, if used carefully and skillfully, deception can be a very effective technique.

As discussed in Chapter 3, the use of deception in performance is not a new idea. In particular, this thesis draws upon the methods of magicians, the professional masters of deception. A magic inspired framework for the use of deception in interaction design and analysis has been presented. This sets out several possible strategies for deception in augmented performance situations, depending on who (or which computer system) is complicit in a deception.

The framework has been used in the design, development and analysis of a set of performances which are presented in this thesis as case studies. These have proven themselves as successful performances in the most real way, in live performance in front of members of the public. The Cup Game in particular demonstrates the relevance of the deception framework to the full development process - from initial system design, to performance development, and in the final application.

As well as being relevant to performance, deception and suspension of disbelief is something that happens in a wide range of situations, meaning that the framework may also be useful in many areas of HCI and also in other areas of computer science such as computer security and software engineering.

Arthur C. Clarke famously stated that "*Any sufficiently advanced technology is indistinguishable from magic.*"[28]. With the use of deceptive techniques to create illusion, it may be possible to reverse this statement; to use magic as a fundamental building block in the creation of advanced performance technology.

Bibliography

Note: For brevity, three regularly referenced conference names have been abbreviated as below:

CHI = *Proceedings of the annual SIGCHI conference on Human factors in computing systems*, ACM, New York

CSCW = *Proceedings of the ACM conference on Computer supported cooperative work*, ACM, New York

DIS = *Proceedings of the conference on Designing Interactive systems*, ACM, New York

Written Work

[1] American Psychological Association. *Code of Conduct*. Washington, DC, USA, 2003.

[2] Anderson, R. Why Information Security is Hard-An Economic Perspective. In *Annual Computer Security Applications Conference*, 2001, page 358, IEEE Computer Society, Washington, DC, USA.

[3] Angeli, A. D., Sutcliffe, A., and Hartmann, J. Interaction, usability and aesthetics: what influences users' preferences? In *DIS*, 2006, pages 271–280.

[4] Aoki, P. M. and Woodruff, A. Making space for stories: ambiguity in the design of personal communication systems. In *CHI*, 2005, pages 181–190.

[5] Bannon, L. and Schmidt, K. CSCW: Four characters in search of a context. In *First European Conference on Computer Supported Cooperative Work*, 1989, pages 358–372, ACM, New York, NY, USA.

[6] Bannon, L. J. From human factors to human actors: the role of psychology and human-computer interaction studies in system design. In *Human-*

computer interaction: toward the year 2000, pages 205–214. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1995.

[7] Bellotti, V., Back, M., Edwards, W. K., Grinter, R. E., Henderson, A., and Lopes, C. Making sense of sensing systems: five questions for designers and researchers. In *CHI*, 2002, pages 415–422.

[8] Benford, S., Crabtree, A., Reeves, S., Sheridan, J., Dix, A., Flintham, M., and Drozd, A. The Frame of the Game: Blurring the Boundary between Fiction and Reality in Mobile Experiences. In *CHI*, 2006, pages 427–436.

[9] Benford, S., Giannachi, G., Koleva, B., and Rodden, T. From Interaction to Trajectories: Designing Coherent Journeys Through User Experiences. In *CHI*, 2009.

[10] Benford, S., Schnädelbach, H., Koleva, B., Anastasi, R., Greenhalgh, C., Rodden, T., Green, J., Ghali, A., Pridmore, T., Gaver, B., Boucher, A., Walker, B., Pennington, S., Schmidt, A., Gellersen, H., and Steed, A. Expected, sensed, and desired: A framework for designing sensing-based interaction. *ACM Trans. Comput.-Hum. Interact.*, 12(1):3–30, 2005.

[11] Berard, F. The Magic Table: Computer Vision Based Augmentation of a Whiteboard for Creative Meetings. *IEEE workshop on Projector-Camera Systems, Nice, France*, 2003.

[12] Billinghurst, M., Kato, H., and Poupyrev, I. The MagicBook - Moving Seamlessly between Reality and Virtuality. *IEEE Comput. Graph. Appl.*, 21(3):6–8, 2001.

[13] Binet, A. Psychology of Prestidigitation. *Annual Report of the Board of Regents of the Smithsonian Institution, 1894*, pages 555–571, 1896.

[14] Birnholtz, J. P., Gutwin, C., and Hawkey, K. Privacy in the open: how attention mediates awareness and privacy in open-plan offices. In *GROUP '07: Proceedings of the 2007 international ACM conference on Supporting group work*, 2007, pages 51–60, ACM, New York, NY, USA.

[15] Bødker, S. When second wave HCI meets third wave challenges. In *NordiCHI '06: Proceedings of the 4th Nordic conference on Human-computer interaction*, 2006, pages 1–8, ACM, New York, NY, USA.

[16] Boehner, K. and Hancock, J. T. Advancing Ambiguity. In *CHI*, 2006, pages 103–106.

[17] Bovermann, T., Groten, J., de Campo, A., and Eckel, G. Juggling Sounds. In *2nd International Workshop on Interactive Sonification, York, UK*, 2007.

[18] Bowers, J. and Hellstrom, S. O. Simple interfaces to complex sound in improvised music. In *CHI*, 2000, pages 125–126.

[19] British Psychological Society. *Code of Conduct*. London, UK, 2006.

[20] Brooks, T. Interpretations: an inter-sensory stimulation concept targeting inclusive access. In *International Conference on Disability, Virtual Reality and Associated Technologies*, 2008, University of Reading, Reading, UK.

[21] Bruegge, B., Teschner, C., Lachenmaier, P., Fenzl, E., Schmidt, D., and Bierbaum, S. Pinocchio: conducting a virtual symphony orchestra. In *ACE '07: Proceedings of the international conference on Advances in computer entertainment technology*, 2007, pages 294–295, ACM, New York, NY, USA.

[22] Buxton, W. Integrating the periphery and context: A new taxonomy of telematics. In *Graphics Interface*, 1995, pages 239–246, Morgan Kaufmann, San Francisco, CA, USA.

[23] Cannell, J. *The Secrets of Houdini*. Hutchinson & Co., London, UK, 1931.

[24] Cao, H., Olivier, P., and Jackson, D. Enhancing Privacy in Public Spaces Through Crossmodal Displays. *Soc. Sci. Comput. Rev.*, 26(1):87–102, 2008.

[25] Card, S. K., Moran, T. P., and Newell, A. The keystroke-level model for user performance time with interactive systems. *Commun. ACM*, 23(7):396–410, 1980.

[26] Castelfranchi, C. Artificial liars: Why computers will (necessarily)deceive us and each other. *Ethics and Inf. Technol.*, 2(2):113–119, 2000.

[27] Christensen, L. Deception in psychological research: When is its use justified? *Personality and Social Psychology Bulletin*, 14(4):664–675, 1988.

[28] Clarke, A. C. *Profiles Of The Future: An Inquiry into the Limits of the Possible*. Phoenix, London, UK, 1962.

[29] Cliff, D. hpDJ: An Automated DJ with floorshow feedback. *HP Labs Tech. Reports*, 2005. `http://www.hpl.hp.com/techreports/2005/HPL-2005-88.html`.

[30] Conan-Doyle, A. *The Edge of the Unknown.* G. P. Putnam, New York, NY, USA, 1930.

[31] Cover, T. and Hart, P. Nearest neighbor pattern classification. *IEEE Trans. on Information Theory*, 13(1):21–27, 1967.

[32] Curzon, P. and McOwan, P. W. Engaging with computer science through magic shows. In *ITiCSE '08: Proceedings of the 13th annual conference on Innovation and technology in computer science education*, 2008, pages 179–183, ACM, New York, NY, USA.

[33] Davey, S. The Possibilities of Mal-Observation. *Proceedings of The Society for Psychical Research*, 4:458–495, 1886.

[34] Davis, S. B., Moar, M., Cox, J., Riddoch, C., Cooke, K., Jacobs, R., Watkins, M., Hull, R., and Melamed, T. 'Ere be dragons: an interactive artwork. In *ACM Multimedia*, 2005, pages 1059–1060, ACM, New York, NY, USA.

[35] DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., and Epstein, J. A. Lying in Everyday Life. *Journal of Personality and Social Psychology*, 70(5):979–995, 1996.

[36] Dessoir, M. The Psychology of Legerdemain. *The Open Court*, 12, 1893.

[37] Dix, A. Challenges for Cooperative Work on the Web: An Analytical Approach. *Computer Supported Cooperative Work*, 6(2-3):135–156, 1997.

[38] Dix, A., Finlay, J., Abowd, G., and Beale, R. *Human Computer Interaction, 3rd Edition.* Prentice Hall, Harlow, UK, 2004.

[39] Dix, A., Sheridan, J. G., Reeves, S., Benford, S., and O'Malley, C. Formalising Performative Interaction. In *12th International Workshop on Design, Specification and Verification of Interactive Systems, Newcastle, UK*, 2005, Springer, Berlin, Germany.

[40] Dourish, P. and Bellotti, V. Awareness and coordination in shared workspaces. In *CSCW*, 1992, pages 107–114.

[41] Dow, S., Mehta, M., Harmon, E., MacIntyre, B., and Mateas, M. Presence and engagement in an interactive drama. In *CHI*, 2007, pages 1475–1484.

[42] Egido, C. Video conferencing as a technology to support group work: a review of its failures. In *CSCW*, 1988, pages 13–24.

[43] Ellis, C. A., Gibbs, S. J., and Rein, G. Groupware: some issues and experiences. *Commun. ACM*, 34(1):39–58, 1991.

[44] Feltham, F., Vetere, F., and Wensveen, S. Designing tangible artefacts for playful interactions and dialogues. In *DPPI '07: Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, 2007, pages 61–75, ACM, New York, NY, USA.

[45] Festinger, L. and Carlsmith, J. M. Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58(2):203–210, 1959.

[46] Fish, R. S., Kraut, R. E., and Chalfonte, B. L. The VideoWindow system in informal communication. In *CSCW*, 1990, pages 1–11.

[47] Fitts, M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6):381–391, 1954.

[48] Fitzkee, D. *Magic by Misdirection*. Lee Jacobs Productions, Pomeroy, OH, USA, 1975.

[49] Gates, C., Subramanian, S., and Gutwin, C. DJs' Perspectives on Interaction and Awareness in Nightclubs. In *DIS*, 2006.

[50] Gaver, W., Beaver, J., and Benford, S. Ambiguity as a resource for design. In *CHI*, 2003, pages 233–240.

[51] Gaver, W., Boucher, A., Law, A., Pennington, S., Bowers, J., Beaver, J., Humble, J., Kerridge, T., Villar, N., and Wilkie, A. Threshold devices: looking out from the home. In *CHI*, 2008, pages 1429–1438.

[52] Gaver, W., Bowers, J., Boucher, A., Law, A., Pennington, S., and Villar, N. The history tablecloth: illuminating domestic activity. In *DIS*, 2006, pages 199–208.

[53] Gaver, W., Sengers, P., Kerridge, T., Kaye, J., and Bowers, J. Enhancing ubiquitous computing with user interpretation: field testing the home health horoscope. In *CHI*, 2007, pages 537–546.

[54] Goffman, E. *Frame Analysis: An Essay on the Organization of Experience.* Harper & Row, London, UK, 1974.

[55] Gutwin, C. and Greenberg, S. Workspace awareness for groupware. In *CHI*, 1996, pages 208–209.

[56] Hancock, J. T., Thom-Santelli, J., and Ritchie, T. Deception and design: the impact of communication technology on lying behavior. In *CHI*, 2004, pages 129–134.

[57] Handel, M. and Herbsleb, J. D. What is chat doing in the workplace? In *CSCW*, 2002, pages 1–10.

[58] Hartmann, J. Assessing the attractiveness of interactive systems. In *CHI*, 2006, pages 1755–1758.

[59] Hergovich, A. The effect of pseudo-psychic demonstrations as dependent on belief in paranormal phenomena and suggestibility. *Personality and Individual Differences*, 36(2):365–380, 2004.

[60] Hertwig, R. and Ortmann, A. Deception in Experiments: Revisiting the Arguments in Its Defense. *Ethics and Behaviour*, 18(1):59–92, 2008.

[61] Hertwig, R. and Ortmann, A. Deception in Social Psychological Experiments: Two Misconceptions and a Research Agenda. *Social Psychology Quarterly*, 71(3):222–227, 2008.

[62] Hinckley, K., Pierce, J., Horvitz, E., and Sinclair, M. Foreground and Background Interaction with Sensor-Enhanced Mobile Devices. *ACM Trans. Comput.-Hum. Interact.*, 12(1):31–52, 2005.

[63] Horn, D. B. Is seeing believing?: detecting deception in technologically mediated communication. In *CHI*, 2001, pages 297–298.

[64] Houdini, H. *A Magician Among The Spirits*. Harper & Row, London, UK, 1924.

[65] Houdini, H. *Houdini on Magic*. Dover Publications, New York, NY, USA, 1953.

[66] Hyman, R. The Psychology of Deception. *Annual Review of Psychology*, 40:133–154, 1989.

[67] Iacucci, G. and Kuutti, K. Everyday Life as a Stage in Creating and Performing Scenarios for Wireless Devices. *Personal Ubiquitous Comput.*, 6(4):299–306, 2002.

[68] Isard, M. and Blake, A. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.

[69] Ishii, H. and Ullmer, B. Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI*, 1997, pages 234–241.

[70] Jakub Segen, J. G., Senthil Kumar. Visual Interface for Conducting Virtual Orchestra. In *ICPR '00: Proceedings of the International Conference on Pattern Recognition*, 2000, page 1276, IEEE Computer Society, Washington, DC, USA.

[71] Jastrow, J. *Fact and Fable in Psychology*. Riverside Press, Cambridge, MA, USA, 1900.

[72] Jorda, S., Kaltenbrunner, M., Geiger, G., and Bencina, R. The reacTable. In *International Computer Music Conference*, 2005, pages 139–146, International Computer Music Association, San Francisco, CA, USA.

[73] Kelley, J. F. An iterative design methodology for user-friendly natural language office information applications. *ACM Trans. Inf. Syst.*, 2(1):26–41, 1984.

[74] Kelman, H. C. Human use of human subjects: the problem of deception in social psychological experiments. *Psychological Bulletin*, 67(1):1–11, 1967.

[75] Kiczales, G., Lamping, J., Mendhekar, A., Maeda, C., Lopes, C., Loingtier, J.-M., and Irwin, J. Aspect Oriented Programming. In *European Conference on Object-Oriented Programming*, 1997, pages 220–242, Springer, Berlin, Germany.

[76] Kim, H. H. J., Gutwin, C., and Subramanian, S. The magic window: lessons from a year in the life of a co-present media space. In *GROUP '07: Proceedings of the 2007 international ACM conference on Supporting group work*, 2007, pages 107–116, ACM, New York, NY, USA.

[77] Kimmel, A. J. In Defense of Deception. *American Psychologist*, 53(7):803–805, 1998.

[78] Koleva, B., Taylor, I., Benford, S., Fraser, M., Greenhalgh, C., Schnädelbach, H., vom Lehn, D., Heath, C., Row-Farr, J., and Adams, M. Orchestrating a mixed reality performance. In *CHI*, 2001, pages 38–45.

[79] Kuhn, G., Amlani, A. A., and Rensink, R. A. Towards a science of magic. *Trends in Cognitive Sciences*, 12(9):323–362, 2008.

[80] Lai, Y.-L. and Hui, K.-L. Internet opt-in and opt-out: investigating the roles of frames, defaults and privacy concerns. In *ACM SIGMIS CPR conference on computer personnel research*, 2006, pages 253–263, ACM, New York, NY, USA.

[81] Lakin, F. A performing medium for working group graphics. In *CSCW*, 1986, pages 255–266.

[82] Laurel, B. *Computers as Theatre*. Addison-Wesley, Boston, MA, USA, 1991.

[83] Levin, G. Computer vision for artists and designers: pedagogic tools and techniques for novice programmers. *AI Soc.*, 20(4):462–482, 2006.

[84] Levisohn, A., Cochrane, J., Gromala, D., and Seo, J. The Meatbook: tangible and visceral interaction. In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, 2007, pages 91–92, ACM, New York, NY, USA.

[85] Love, S. and Perry, M. Dealing with mobile conversations in public places: some implications for the design of socially intrusive technologies. In *CHI*, 2004, pages 1195–1198.

[86] Lyons, M. J. and Tetsutani, N. Facing the music: a facial action controlled musical interface. In *CHI*, 2001, pages 309–310.

[87] Marshall, J., Benford, S., and Pridmore, T. Eye-balls: juggling with the virtual. In *ACM Creativity & Cognition*, 2007, pages 265–266, ACM, New York, NY, USA.

[88] Marshall, J., Benford, S., and Pridmore, T. Eye-Balls: Cheap and Cheerful Interactive Performance. *Leonardo*, 41(3):304–305, 2008.

[89] Marshall, J., Mills, S., and Benford, S. Using Object Interactions to Improve Particle Filter Performance. In *British Machine Vision Conference*, 2006, pages 337–346, British Machine Vision Association, Cambridge, UK.

[90] Marshall, J., Pridmore, T., Pound, M., Benford, S., and Koleva, B. Pressing the Flesh: Detecting Multiple Touch and Finger Pressure on Arbitrary Surfaces. In *Proceedings of Pervasive 2008*, 2008, pages 38–55, Springer, Berlin, Germany.

[91] Marti, P., Pollini, A., Rullo, A., and Shibata, T. Engaging with artificial pets. In *EACE '05: Proceedings of the 2005 annual conference of the European association of cognitive ergonomics*, 2005, pages 99–106, University of Athens, Athens, Greece.

[92] Mathew, A. P. and Taylor, J. Auralscapes: engaging ludic ambiguity in the design of a spatial system. In *CHI*, 2008, pages 2533–2542.

[93] Maurer, D. *The Big Con*. Reprinted 2000 - Arrow Books, London, UK, 1940.

[94] Milgram, S. Behavioral Study of Obedience. *Journal of Abnormal and Social Psychology*, 67(4):371–378, 1963.

[95] Mitnick, K. D. *The Art of Deception*. John Wiley & Sons, Chichester, UK, 2002.

[96] Moraveji, N., Li, J., Ding, J., O'Kelley, P., and Woolf, S. Comicboarding: using comics as proxies for participatory design with children. In *CHI*, 2007, pages 1371–1374.

[97] Mulder, M. and Stemerding, A. Threat, attraction to group, and need for strong leadership. *Human Relations*, 16(4):317–334, 1963.

[98] Murphy, D., Andersen, T. H., and Jensen, K. Conducting audio files via computer vision. In *5th International Workshop on Gesture and Sign Language based Human-Computer Interaction*, 2003.

[99] Nardi, P. Toward a Social Psychology of Entertainment Magic (Conjuring). *Symbolic Interaction*, 7(1):25–42, 1984.

[100] Neagoe, V. and Bishop, M. Inconsistency in deception for defense. In *NSPW '06: Proceedings of the 2006 workshop on New security paradigms, Schloss-Dagstuhl, Germany*, 2006, pages 31–38, ACM, New York, NY, USA.

[101] Norman, D. A. *The Design of Everyday Things*. MIT Press, Cambridge, MA, USA, 1990.

[102] Notes and Queries. Inspector Sands Question. *Guardian Newspaper*. `http://www.guardian.co.uk/notesandqueries/query/0,5753,-22390,00.html`.

[103] Quittner, J. Penn. *Wired*, 2(9), 1994.

[104] Raento, M., Oulasvirta, A., Petit, R., and Toivonen, H. ContextPhone - A prototyping platform for context-aware mobile applications. *IEEE Pervasive Computing*, 4(2):51–59, 2005.

[105] Rampton, J. Kidnapped! And All in the Name of Art. *The Independent, 3rd June*, 1998.

[106] Randi, J. The Psychology of Conjuring. *Technology Review*, 80(3):56–63, 1978.

[107] Randi, J. *An Encyclopedia of Claims, Frauds, and Hoaxes of the Occult and Supernatural*. St Martin's Press, New York, NY, USA, 1997.

[108] Raphael Grasset, M. B., Eric Woods. Art and Mixed Reality: New Technology for Seamless Merging Between Virtual and Real. *Media-Space Journal*, 1(1), 2007.

[109] Reeves, S., Benford, S., O'Malley, C., and Fraser, M. Designing the spectator experience. In *CHI*, 2005, pages 741–750.

[110] Reynolds, M., Schoner, B., Richards, J., Dobson, K., and Gershenfeld, N. An immersive, multi-user, musical stage environment. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, pages 553–560, ACM, New York, NY, USA.

[111] Robert-Houdin. *The Secrets of Stage Conjuring*. Reprinted 2008 - Wildside Press, Rockville, MD, USA, 1881.

[112] Rowe, N. C. Designing Good Deceptions in Defense of Information Systems. In *ACSAC '04: Proceedings of the 20th Annual Computer Security Applications Conference*, 2004, pages 418–427, IEEE Computer Society, Washington, DC, USA.

[113] Said, N. S. An engaging multimedia design model. In *IDC '04: Proceedings of the 2004 conference on Interaction design and children*, 2004, pages 169–172, ACM, New York, NY, USA.

[114] Satchell, C. Cultural theory and real world design: Dystopian and Utopian Outcomes. In *CHI*, 2008, pages 1593–1602.

[115] Schnadelbach, H., Penn, A., Steadman, P., Benford, S., Koleva, B., and Rodden, T. Moving office: inhabiting a dynamic building. In *CSCW*, 2006, pages 313–322.

[116] Sengers, P. and Gaver, B. Staying open to interpretation: engaging multiple meanings in design and evaluation. In *DIS*, 2006, pages 99–108.

[117] Sheridan, J., Bayliss, A., and Bryan-Kinns, N. iPoi: acceleration as a medium for digital live art. In *Proceedings of UbiComp 2006*, 2006, Springer, Berlin, Germany.

[118] Sheridan, J. G., Bryan-Kinns, N., and Bayliss, A. Encouraging Witting Participation and Performance in Digital Live Art. In *21st British HCI Group Annual Conference*, 2007, British Computer Society, Swindon, UK.

[119] Sheridan, J. G., Dix, A., Lock, S., and Bayliss, A. Understanding Interaction in Ubiquitous Guerrilla Performances in Playful Arenas. In *Proceedings of HCI 2004*, 2004, pages 3–18, British Computer Society, Swindon, UK.

[120] Simons, D. J. and Chabris, C. F. Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, 28(9):1059–1074, 1999.

[121] Sparacino, F., Davenport, G., and Pentland, A. Media in performance: Interactive spaces for dance, theater, circus, and museum exhibits. *IBM Systems Journal*, 39(3 + 4), 2000. `http://alumni.media.mit.edu/~flavia/DanceSpace/dancespace.html`.

[122] Steinberg, N. Dude, Where's my Video Phone. *Forbes.com: Special Report - The Future*, 2007. `http://www.forbes.com/2007/10/13/neil-steinberg-videophone-tech-future07-cx_ns_1015gizmo.html`.

[123] Stewart, J., Bederson, B. B., and Druin, A. Single display groupware: a model for co-present collaboration. In *CHI*, 1999, pages 286–293.

[124] Stoll, C. Stalking the wily hacker. *Commun. ACM*, 31(5):484–497, 1988.

[125] Sun, Y., Hollerbach, J., and Mascaro, S. Imaging the Finger Force Direction. In *IEEE Conference on Computer Vision and Pattern Recognition*, June 2007, pages 1–6, IEEE Computer Society, Washington, DC, USA.

[126] Tang, A., Finke, M., Blackstock, M., Leung, R., Deutscher, M., and Lea, R. Designing for bystanders: reflections on building a public digital forum. In *CHI*, 2008, pages 879–882.

[127] Tatler, B. W. and Kuhn, G. Don't look now: The magic of misdirection. In van Gompel, R., Fischer, M., Murray, W., and Hill, R., editors, *Eye Movement Research: Insights into mind and brain*, pages 697–714. Elsevier, Amsterdam, Netherlands, 2007.

[128] Tognazzini, B. Principles, techniques, and ethics of stage magic and their application to human interface design. In *CHI*, 1993, pages 355–362.

[129] Triplett, N. The Psychology of Conjuring Deceptions. *American Journal of Psychology*, 11(4):493–510, 1900.

[130] vom Lehn, D., Hindmarsh, J., Luff, P., and Heath, C. Engaging constable: revealing art with new technology. In *CHI*, 2007, pages 1485–1494.

[131] Wallace, A. R. The Late Mr S.J. Davey's Experiments. *Journal of The Society for Psychical Research*, 6:34–37, 1893.

[132] Willier, A. and Marque, C. Juggling gestures analysis for music control. *Lecture Notes In Computer Science*, 2298:296–306, 2001.

[133] Wiseman, R. and Greening, E. 'It's still bending': Verbal suggestion and alleged psychokinetic ability. *British Journal of Psychology*, 96(1):115–127, 2005.

[134] Wright, P. and McCarthy, J. Empathy and experience in HCI. In *CHI*, 2008, pages 637–646.

Artworks, games etc.

[135] Active Ingredient. *Heartlands*. Performance/Game, 2007.

[136] Aerotech. *GloBalls demonstration videos*. Web Video, 2006. `http://www.globall.com`.

[137] Alicia. *Webcam "Magic" Trick*. Web Video, 2008. `http://uk.youtube.com/watch?v=BrMhmz6jYO8`.

[138] Anonymous. *DDR with hands*. Web Video, 2008. `http://uk.youtube.com/watch?v=8gOXOkpTYvU&feature=related`.

[139] Atari Games. *Gauntlet*. Software, 1985.

[140] Bishop, D. *Marble Answering Machine*. Design, 2002.

[141] Blast Theory. *Kidnap*. Performance, 1998.

[142] Bumgardner, J. *iPolygraph*. Web Based Trick, 2008. `http://krazydad.com/ipolygraph/secret.php`.

[143] Dorner, W. *Anywhere, Somewhere, Everywhere*. Interactive Tour, 2007.

[144] Iaconesi, S. and Persico, O. *Degradazione per Sovrapposizione di Corpi*. Installation, 2008.

[145] Jay, J. *High Tech Juggling Jukebox*. Performance, 1994. `http://jamesjay.com/juggling/jukebox/hightech/index.html`.

[146] Kreuger, M. *Videoplace*. Installation, 1969.

[147] Levin, G. and Lieberman, Z. *Messa di Voce*. Performance, 2003.

[148] Lewis, T. *Pony*. Installation, 2008.

[149] Litchfield, J. *Mother of All Online Card Tricks*. Web Based Trick, 2008. `http://jason.ushospitality.com/diamonds/reflectedcard38.swf`.

[150] Lozano-Hemmer, R. *Frequency and Volume*. Installation, 2008.

[151] Maita, A. *Bandai Tamagotchi*. Toy, 1996.

[152] Malia, J. *PRIVATE PUBLIC*, 2004.

[153] Milne, A. A. *Winnie the Pooh*. Methuen & Co. Ltd., London, 1926.

[154] Nauman, B. *Live/Taped Video Corridor*. Installation, 1970.

[155] Nintendo. *The Legend of Zelda: Phantom Hourglass*. Software, 2007.

[156] Palindrome. *EyeCon*. Computer System, 1999. `http://eyecon.palindrome.de/`.

[157] Rokeby, D. *Very Nervous System*. Computer System, 1986.

[158] Salom, J. Performance. `http://www.joelsalom.com`.

[159] Sorayama, H. *Sony AIBO*. Toy, 1999.

[160] Stelarc. *Ping Body*. Performance, 1996.

[161] Tiger Electronics. *Furby*. Toy, 1995.

[162] Ubisoft. *Petz*. Software, 1995.

[163] Young, M. *SuperTed in Space*. Frederick Muller, London, 1980.

Appendix A

Using Object Interactions to Improve Particle Filter Performance

Joe Marshall, Steve Mills, Steve Benford,
Presented at BMVC 2006, Edinburgh, UK
{jqm,smx,sdb}@cs.nott.ac.uk

This paper describes and evaluates a novel set of approaches to handle situations where multiple distinct and visually differing objects are tracked, such as tracking of people and objects they are manipulating. Unlike tracking of multiple similar objects, visually different interacting objects can provide an opportunity to improve the tracking accuracy. These approaches are designed for use with Condensation/Particle Filter based algorithms, and allow drop-in replacement of tracker modules for each object type tracked. They use information about the relationships and interactions between objects to improve the tracking, rather than in order to distinguish between the objects, as in current algorithms. They are also designed to be highly efficient, for real time use. The approaches are tested on a challenging set of real data and achieve tracking performance similar to using a single very high dimensional tracker, but with vastly reduced complexity and hence much better time performance.

A.1 Introduction

The Condensation[4] algorithm is commonly used for single object tracking in video sequences. There are many generalizations of the algorithm to cope with the tracking of multiple similar objects[3, 5, 6]. These attempt to disambiguate the multiple objects, to allow association between each particular object over multiple frame sequences and to maintain tracking of the full set of objects. This paper addresses the class of situations where multiple distinct and visually differing objects are tracked at the same time, such as playesrs balls and rackets etc. in games, or tracking of people manipulating objects. In these situations, interaction between objects, rather than being a problem, creates an opportunity to improve the tracking accuracy for the individual objects.

There are two main ways of extending single object tracking to tracking multiple objects, the first is to use a single object tracker which tracks over a multi-dimensional space representing the combined state of all objects tracked. This approach has serious performance problems, due to the high dimensionality of the tracking space. Ways to overcome this have been suggested[2], but these are still complex and performance intensive.

The second method of tracking multiple objects is to use a single tracker per object, and to add some way of taking the dependencies between the objects into account. This paper investigates various methods for taking these dependencies into account, without significantly increasing the complexity of the multiple object tracking process, in order to develop real-time suitable tracking methods for tracking multiple interacting objects. It is possible to make use of the difference between objects in a way that current multiple tracking algorithms are not able to, thus allowing for a more efficient and accurate tracking process.

A.1.1 Existing Work

The Condensation algorithm is commonly used to detect object position from video data[4]. It uses a large number of particles each storing a single hypothesis about the possible state of the object, and an observation of the input data.

The model works in four stages. Firstly, each particle is scored as to how well it fits the observation data. From this set of scored particles, a new set of particles with the same number of elements as the current set is generated. This is generated by weighted random sampling, based on the score of each particle, so some high scoring hypotheses may be chosen multiple times and some hypotheses may not be propagated into the new set at all. After this sampling stage, the new set of items is subjected to drift, which uses a motion model to predict how this particle will have moved since the last frame, and diffusion, which adds a certain amount of randomness. This predicts a new position for the particle. Finally, in the measurement phase, each particle in the new set is scored against the video frame. An estimate of the current position of the object is derived from this set, usually by using a weighted mean.

There are several extensions of Condensation to multiple tracking of similar objects which effectively use a single tracker per object tracked. Khan et al.[5] use Markov Random Fields to model the interaction between several ants they are tracking. At points when the ants are far away from each other, their method works as a set of standard particle filters. Similarly, various methods, eg.[3, 9, 10], use one multi-modal particle distribution for all objects. These essentially try to solve the data association problem of how to assign observations to individual objects and to maintain the modality of the particle distribution when objects interact. These methods are not generally useful when objects are visually distinct, as the association problem does not exist.

Han et al.[2] represent the state of two tracked objects within one high dimensional particle, but store Gaussians representing the modes in the distribution, rather than individual point particles. This reduces the number of particles required for tracking in high dimensional spaces. However, while this increases efficiency, it is still exponential in the number of objects, and only increases the number of dimensions it is possible to track, rather than avoiding the underlying problems of using a high dimensional state to track multiple objects. This means it still requires large numbers of particles in order to track complex systems with more than two interacting objects. It is also inefficient at points when objects are not currently interacting, and could be tracked equally well individually. Wu et al.'s co-inference tracking[11] uses multiple cues in tracking a single object, however this does not simply extend to use multiple different objects as cues, as the cues do not always reinforce each other and the method fails.

A.1.2 The Example Tracking System

Situations where this kind of tracking occurs include tracking players and balls in sport[12], and tracking people manipulating objects during stroke rehabilitation[1]. A system tracking a person juggling is used here as an algorithm test-bed. Juggling, provides a challenging tracking situation, with a lot of interaction between the fast moving balls and hands, but is easy to set up and control. This uses three different trackers, all of which work based on the detection of particular colours and shapes in the image. The face tracker keeps track of the position of the persons face in order to determine where the centre-line between the two shoulders is located. The arm tracker detects two arms, on either side of this line. Finally the ball tracker detects the number, positions and velocity of multiple balls and tries to detect whether they are currently held by the juggler, or are in the air.

A.2 The Problem

It would be ideal to use a single Condensation tracker in which at time t each particle holds the combined state of a set of K objects $Object_{1...K}$. This tracker uses a scoring function to calculate a probability for a combined state $X_t = \{X_t^1, X_t^2..., X_t^K\}$, based on an observation Z_t, i.e. $P(X_t^1 \cap X_t^2 ... \cap X_t^K | Z_t)$.

It is desirable to approximate this single tracker with a set of individual trackers, with state vectors $X^1, ..., X^K$. The state space for each tracker will have a significantly lower number of dimensions than the single tracker and hence will require smaller numbers of particles to achieve the same tracking performance.

However, the objects are not independent, so using completely independent trackers, without tracking the interactions between the objects, would fail to exploit the available information. Given this fact, the probabilities being calculated for tracker k must approximate $P\left(X_t^k|X_t^1 \cap X_t^2...X_t^{k-1} \cap X_t^{k+1}..X_t^K, Z_t\right)$. It is not possible to directly calculate these conditional probabilities for each particle, as this would involve effectively calculating a high-dimensional tracker.

A.2.1 The Proposed Solution

The interactions affecting object X^1, based on $X^{2..K}$ are considered here, as the algorithm is identical for other objects.

A.2.1.1 Definitions

$\underline{X}_t^{1..K}$	This is a set of random variables that represent the part of the X^k which has an effect on the other X states. They are based on the value of each of the X states from the previous frame by using the prediction function inherent in the particle filter. $\underline{X}_t^k$ is only dependent on X_{t-1}^k, and is assumed conditionally independent of the observation and the other $\underline{X}$ variables. They are parameterized in a multi-dimensional space, which is discrete, or may be discretised.
$f_{1..K}(x)$	This is the probability density function of $\underline{X}_t^k|X_{t-1}^k$, and must be computationally simple to calculate given a particular X^k state.
$E^k(X^k)$	This is the region of the $\underline{X}^k$ space where a given X^k has an effect. i.e. $x \notin E^k(X^k) \implies P(\underline{X}^k = x|X^k) = 0$, for a given X^k
$R^{1,k}(X^1)$	$P(X^1|\underline{X}^k)$ is dependent on only the region of $\underline{X}^k$ space defined by this function, i.e. for any given X^1, if $x \notin R^{1,k}(X^1)$ then $P(X^1|\underline{X}^k = x) = P(X^1|X^1 \notin R^{1,k}(X^1))$, which is constant for that X^1
T^k	This is an overall region containing all points where $P(\underline{X}_t^k|X^k)$ may be non zero for all X^k. i.e. it is the union of all possible E^k regions.

A.2.1.2 Algorithm

Taking the ideal tracker, with particle score equal to $P(X_t^1 \cap X_t^2... \cap X_t^K|Z_t) = P(X^2...\cap X^K|Z)P(X_t^1|X_t^2...\cap X_t^K, Z_t)$, it is assumed that the states of the variables $X^{2..K}$for the previous frame are known, and the score $P(X_t^1|X^2... \cap X^K, Z)$ is to be calculated. This score can be rewritten in terms of $\underline{X}_t^{2..K}$, and as these are independent, in terms of $f_{1..K}$:

$$P(X_t^1|X_t^2...\cap X_t^K, Z) \approx \left[\sum_{\{x_2\in T_2,...x_K\in T_K\}} \Big(f_2(x_2)...f_K(x_K)P(X_t^1|\underline{X}_t^2 = x_2...\cap \underline{X}_t^K = x_K, Z)\Big)\right]. \tag{A.1}$$

For two object interaction, this calculation is simple enough, however, for multiple object interaction, further simplification is required to avoid the combinatorial nature of this equation. An assumption is made that $P(X^1|X_t^2,...X_t^K)$ can be aproximated using a function of the pairwise interactions. For example, in the test tracker, the interaction function:

$$P(A|F, B_1, ..., B_C, Z) \approx [P(A|F,Z)]\left[\sum_{1..C}\frac{P(A|B_c,Z)}{C}\right], \tag{A.2}$$

where A=arm position,F=face position,B_c=cth ball position(of C balls), is used to score the arm. To use this approximation, Equation A.1 is calculated for each two object case included in the equation. This approximation in terms of pairwise interactions has been shown to be relatively efficient in testing. For the rest of this section only pairwise interactions are considered, between X^1 and X^2.

By definition, $P(X^1|\underline{X}^2)$ only needs to be evaluated within $R^{1,2}$ so

$$P(X^1|X^2,Z) \approx \sum_{x_2\in R^{1,2}} \Big(f_2(x_2)P(X^1|\underline{X}^2 = x_2, Z)\Big) + \sum_{x_2\notin R^{1,2}} \Big(f_2(x_2)P(X^1|\underline{X}^2 \notin R^{1,2}, Z)\Big). \tag{A.3}$$

As $P(X^1|\underline{X}^2 \notin R^{1,2}, Z)$ is not dependent on x_2, and thus can be taken outside the sum, and f_2 must sum to one, the dependency on values of f_2 outside of $R^{1,2}$ may be removed, giving

$$P(X^1|X^2,Z) \approx \sum_{x_2\in R^{1,2}} \Big(f_2(x_2)P(X^1|\underline{X}^2 = x_2, Z)\Big) + P(X^1|\underline{X}^2 \notin R^{1,2}, Z)\left(1 - \sum_{x_2\in R^{1,2}} f_2(x_2)\right). \tag{A.4}$$

This is used as the input to the interaction function. Similar inputs are created for other interactions between the objects. In many cases, certain objects will be independent of others, in which case they will not be in the interaction function, and hence not require calculation. For example, in the test system the face tracker is independent, so does not require any extra inputs to be calculated in its interaction function. However, this does not directly allow the creation of a set of trackers as the $f_{1...K}$ pdfs rely on the state of individual particles, an approximation to the distribution of the $f_{2..K}$ values must be used as input to the X^1 tracker, giving a two step process:

1. Calculate probabilities for all $X^{1..K}$ particles based on the observation data and $f_{1..K}$ generated from the previous frame.

2. Calculate next frame $P(\underline{X}^k|X^k)$ and the pdfs f_k for all particles and construct approximations to the distributions of the pdfs, weighted by particle probability.

A.2.2 Calculating the Approximate Distribution

It is important that an efficient way of calculating and storing these distributions be used. Otherwise, no efficiency will be gained over the higher dimensional calculation. Three ways of calculating and storing the intermediate distributions of $f_{1..K}$ are described here.

A.2.2.1 Gaussian Mixtures

This method attempts to approximate the overall state of the $f_{1..K}$ functions by using a set of N Gaussians for each pdf, representing the mean values of $f_{1..K}$ at each position, weighted by the particle score. First the generated probabilities $P(\underline{X}^k|X^k)$ are partitioned spatially, then their contribution to a single Gaussian for each partition can be estimated using an unbiased Maximum Likelihood Estimator. In the example system it is possible to generate these Gaussians directly from the X^k states and $P(X^k|Z)$ values without actually calculating all the f_k values at any point.

Samples from these Gaussians are used as input to the scoring function instead of sampling f_n. The sum of all Gaussians at position x is used as an estimate of the pdf at that position, and hence a pairwise score function is created as:

$$\sum_{x \in R^{1,2}} \left(P(X^1|\underline{X}^2 = x, Z) \sum_{n \in 1..N} [G_n(x)] \right) + P(X^1|\underline{X}^2 \notin R^{1,2}, Z) \left(1 - \sum_{x \in R^{1,2}} \sum_{n \in 1..N} [G_n(x)] \right). \tag{A.5}$$

The information from the Gaussians as to the distribution means that the calculation of this sum may be simplified to avoid having to calculate a large number of Gaussian values, by either using a single Gaussian sample as an estimate, or using the Gaussian distribution parameters directly. In this way, it is possible to avoid calculating many values from the Gaussians, making this technique relatively efficient.

This technique has advantages, in that the Gaussians are easily calculated and can be used directly in the second stage. However, it reduces the usefulness of the Condensation algorithm itself, as the modality of the data is reduced to the number of Gaussians used in the partitioning stage. It may be possible to detect when this is occurring, by detecting that a Gaussian with a large variance is being created, and thus split a partition automatically, however, this did not seem reliable in testing.

A.2.2.2 Sampling

For each X^1 state, a state from X^2, or a small set of states is sampled, to use as input to the X^1 state. This is done by using a probabilistic sampling stage similar to the Condensation sampling stage. The test system makes use of the Condensation sampling stage, by sampling uniformly from the distribution created by the sampling stage of the Condensation algorithm, thus avoiding doing two similar sampling stages. This method is appealing in that it uses a method very similar to the Condensation algorithm, and potentially allows for a full representation of the modality of the data. It is also the most simple conceptually of the methods described. A set of 10 states was used for the test version of this, as using higher numbers of states was too slow for real time use.

A.2.2.3 Mean Probability Density Functions

This combines aspects of the two methods above, to create a technique that allows multi-modality, but also allows all X^2 states to influence the result, unlike the sampling method. Instead of partitioning the particles as in the Gaussian model, a uniform partitioning over a fixed grid is used. At each point, the weighted mean of all the f_2 values is calculated:

$$M[x] = \frac{\sum_{1..count(particles)} P(X^2_{t-1}[particle]|Z_{t-1})P(\underline{X}^2_t = x|X^2_{t-1}[particle])}{\sum_{1..count(particles)} P(X^2_{t-1}[particle]|Z)}. \quad \text{(A.6)}$$

The X tracker can then use $M[x]$ as input to its scoring function, defined as

$$P(X^1|X^2, Z) \approx \sum_{x \in R^{1,2}} P(X^1|\underline{X}^2 = x, Z)M[x] + P(X^1|\underline{X}^2 \notin R^{1,2}, Z)\left(1 - \sum_{x \in R^{1,2}} M[x]\right). \quad \text{(A.7)}$$

The values of $M[x]$ are the same over all particles, thus given that $P(\underline{X}|X) > 0$ only within a fixed range, all possible values of $M[x]$ can be precalculated. An extra stage in the Condensation algorithm is used to calculate M. This can be done by using a matrix to accumulate probability values for each X particle. For this to work efficiently, $E^2(X^2)$ and $R^{1,2}(X^1)$ must be relatively localised for states X^1 and X^2, so that each particle does not have to add to or read many accumulators. The process for each tracker is now:

1. Calculate tracker, using $M[x]$ matrices in scoring function.
2. Calculate the $M[x]$ matrix based on X^k distribution.

This is slightly less time efficient than using a small number of Gaussian mixtures. However, it can more accurately represent the distribution of the particles, thus making it work much better in situations where tracker distribution is noisy and the tracker is tracking several modes in the observation data.

A.2.3 Extensions

In the worst case scenario, all K objects will interact with each other. This will require K intermediate representations to be written, one per object, and $K-1$ to be read in each scoring function. There will still be a vastly smaller total number of particles compared to the single tracker, and thus it will be far more efficient, especially for large K. However, in many real cases optimisations will be possible where some objects are not affected by others, such as in our head tracker, which does not take input from any other trackers. It is also useful to use methods such as mixture tracking[10] within individual trackers, where there are multiple objects of one type. If it is unlikely that the multiple similar objects will overlap, or it is not relevant to the other scoring functions, a combined intermediate representation, representing the sum of the object pdfs may be used, rather than using a separate pdf for each object. This is the case in the ball and arm trackers described below.

In the case of juggling, the system is tracking balls, which are relatively predictable except when interacting with the hands, and hands, which are much less predictable. It is useful to reduce the reliance on the next frame predictions as much as possible for the less predictable object. In this case, the output from the less predictable tracker is used directly to create the probability density function, rather than using the next frame prediction. This is likely to be useful in other similar situations where a human interacts with one or more relatively predictable objects.

A.3 Evaluation

A.3.1 Implementation Details

Only the mean pdf based tracker is described in depth here, the implementation was similar for other tracking methods. The juggling system makes extensive use of tracker interaction as shown in Figure A.1.

Face Tracker This has particles defining face position, size and angle of an ellipse representing the face shape. Each particle's score is added to a pdf for each position in the face ellipse, and normalized to create a representation of the distribution of the likely face position.

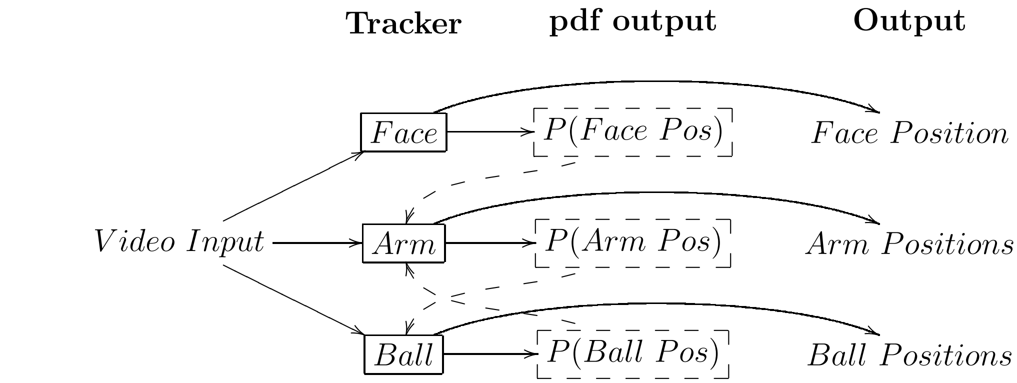
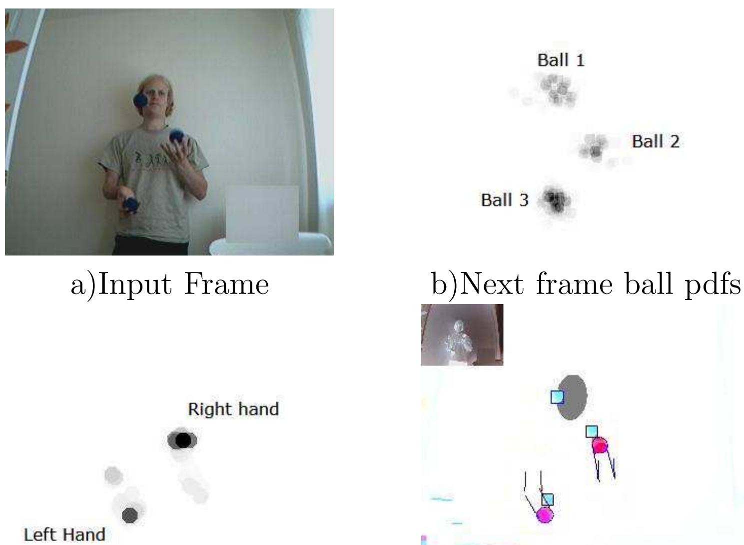


Figure A.1: Use of pdfs in the juggling tracker



a)Input Frame b)Next frame ball pdfs

c)Hand Position pdfs d)Output - Arm and Ball tracking

Figure A.2: The tracker state for a single frame

Arm Tracker When calculating a score for a particular arm position, this tracker disregards areas of skin colour which are part of the face by using the face probability map. Prior to this, the tracker was often misled by the face pixels. The arm tracker also uses a ball position probability map generated by the ball tracker in the previous frame. This is because when a ball is stopped over, or heading towards the arm, it generally defines the position of the hand. The arm tracker outputs a combined pdf for both hands, giving the sum of the probabilities for the left or right hand being centred on a particular position, as in Figure A.2c. (annotated to show the modes of the left and right hand).

Ball Tracker This uses the pdfs generated by the arm tracker to detect caught balls. This is both because the 'caught/not caught' status is a desired output, and because caught balls are often partly occluded by the hand, and may otherwise be lost by the ball tracker, or inaccurately tracked. The position and velocity of the ball is used to generate a combined pdf of the position of all balls in the next frame for the arm tracker as in Figure A.2b. Multiple balls are detected by use of a segmentation algorithm on the output of the ball tracker, this segmentation is also used to ensure that one ball does not take over all the particles, in similar way to Milstein et al.[8]

A.3.2 Testing Results

Testing was run using a Pentium 4 2.6GHz test machine. Up to 4000 particles in total were used (3000 ball, 900 arm, 100 face), this number of particles allowed processing of the video at approx 30 frames per second, on all interaction methods.

In order to compare against the 'ideal' tracker (as described in section A.2) a single 20 dimensional particle filter was used. This was only able to track 1 ball (plus two arms and a face), due to the vast number of particles required for multiple balls being too large to fit in memory/disk space. This took several minutes per frame to run. A set of independent particle filters, using only observation information were also used.

Three test sequences (available at http://www.mrl.nott.ac.uk/~jqm/juggling/bmvc) were used, firstly a 500 frame sequence (plus 100 frames of initialisation), of 1 ball being thrown from hand to hand, involving various complicated hand movements and different throws, secondly a 1400 frame sequence of mainly 3 ball juggling, with a short section of 2 and 1 ball, and finally a set of 200 frame sequences of juggling up to 4 balls.

In this tracker, the desired output is whether a ball is caught, and its position. When the balls are in flight and not nearing catch or release, all methods provide good tracking performance. In situations where the ball is being caught or held, there were several possible errors that occurred. These are defined as minor or serious, depending on their effect on the data being collected. The tracker is able to automatically reinitialize after an error and detect new objects, (it uses a small number of initialization particles, as in [7]), so no manual reinitialization was required after an error.

Serious Errors A ball is lost and the tracker has to reinitialize. The tracker detects a catch when no catch has occurred. A catch occurs and the tracker doesn't notice it.

Minor Errors A ball is temporarily lost but re-tracked without reinitialization, which may occur in some occlusion situations, and is not generally a problem for the system. An arm fails to be tracked - this may cause other errors to occur, but usually the next time a ball interacts with the missing arm, the tracker recovers.

Tracker Type	Particles	1 ball errors		3 Ball Errors	
		Minor	Serious	Minor	Serious
Ideal	120,000	4	0	n/a	n/a
	1,000,000	2	0	n/a	n/a
Independent	1,000	4	22	11	97
	4,000	1	28	15	56
Gaussian	1,000	5	4	8	13
	4,000	4	2	10	8
Sampling	1,000	8	2	14	11
	4,000	5	1	15	7
Mean pdfs	1,000	2	4	10	7
	4,000	2	0	5	4

Table A.1: Comparative Tracker Performance

Using the one ball sequence, it was possible to compare the results from the ideal tracker against the methods described in section A.2.2. For the 3 ball sequence, only the relative performance of the different types of tracker interaction could be compared. The testing results are shown in Table A.1. In the 1 ball test 1,000,000 'ideal' particles were required in order to achieve the power of 4000 particles using the best performing mean pdf method. The independent trackers were very hard to tune in order to get a balance between false catches and missed catches, hence the large number of serious errors shown, which are mainly catch errors. Early attempts to fix these issues led to the algorithm presented here.

The best performing, pdf based tracker was also tested on the 200 frame sequences. For these sequences a ground truth hand position for one hand was marked and the distance from this point to the tracker hand circle was measured for each frame. Figure A.3, shows for each number of balls, the percentage of frames within each error value, and the mean error. This was interesting, as with no balls, the arm tracker was very poor. Once balls were being juggled, the tracking accuracy went up greatly; also slightly improved tracking was observed as higher numbers of balls were juggled. These results demonstrate how the individual trackers are made more robust using feedback between the trackers.

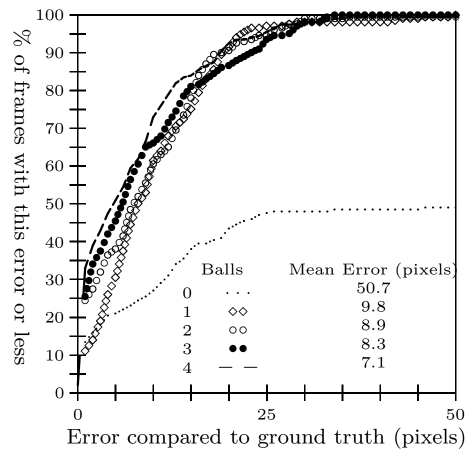


Figure A.3: pdf based Arm tracker accuracy for different numbers of balls

A.4 Conclusions

The results shown in this paper demonstrate that these methods can greatly improve the tracking accuracy of individual object trackers. The final accuracy is similar to that of a single high-dimensional tracker, and far better than independent trackers. The proposed methods do not significantly increase the complexity, and hence perform only slightly slower compared to completely independent trackers, allowing real time use of these methods. This paper shows that where there are multiple visually distinct objects in a scene, this is no data association problem. Rather, there is an opportunity to improve the tracking performance for each individual object, to make it better than that of independent trackers. In addition, the improved tracking due to the interactions allows the use of simpler, faster individual trackers, as demonstrated by the very simple arm tracker, which performs poorly standalone, but has much better performance when interactions are taken into account.

These methods also offer a major advantage over existing algorithms in that they allow the use of individual trackers as black boxes, so that trackers for one class of object can easily be modified without needing to alter other trackers. For example in this system, several different versions of the arm and face trackers have been used, which may just be swapped in, without altering the ball tracker. They are also highly generalisable, and can be used in many situations where multiple trackers are used, simply by altering measurement functions and creating functions to generate the intermediate probabilities.

Bibliography

[1] A. Ghali, A. Cunningham, and T. Pridmore. Object and event recognition for stroke rehabilitation. In *VCIP*, 2003.

[2] B. Han, D. Comaniciu, Y. Zhu, and L. Davis. Incremental density approximation and kernel-based Bayesian filtering for object tracking. In *CVPR*, 2004.

[3] C. Hue, J.-P. L. Cadre, and P. Perez. A particle filter to track multiple objects. *IEEE Workshop on Multi-Object Tracking*, pages 61–68, 2001.

[4] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.

[5] Z. Khan, T. Balch, and F. Dellaert. MCMC-based particle filtering for tracking a variable number of interacting targets. *PAMI*, 27(11):1805–1918, 2005.

[6] E. B. Koller-Meier and F. Ade. Tracking multiple objects using the Condensation algorithm. *Journal of Robotics and Autonomous Systems*, pages 93–105, 2001.

[7] E. B. Meier and F. Ade. Using the Condensation algorithm to implement tracking for mobile robots. In *Third European Workshop on Advanced Mobile Robots*, 1999.

[8] A. Milstein, J. N. Sánchez, and E. T. Williamson. Robust global localization using clustered particle filtering. In *AAAI/IAAI*, pages 581–586, 2002.

[9] D. Tweed and A. Calway. Tracking many objects using subordinated Condensation. In *BMVC*, pages 283–292, 2002.

[10] J. Vermaak, A. Doucet, and P. Pérez. Maintaining multi-modality through mixture tracking. In *ICCV*, pages 1110–1116, 2003.

[11] Y. Wu and T. S. Huang. A co-inference approach to robust visual tracking. In *ICCV*, 2001.

[12] F. Yan, W. Christmas, and J. Kittler. A tennis ball tracking algorithm for automatic annotation of tennis match. In *BMVC*, 2005.

Appendix B

Cup Game Script

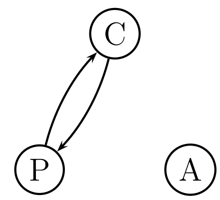
This appendix shows the script that is used by the performer in the Cup Game in Chapter 5. It is formatted in two columns. The left column gives the script and associated directions itself. The right column describes the deceptions occuring in each section, and the effect that this may have on the audience.

The script is presented as it was used in the Cup Game, although as in any performance of this nature, some level of improvisation and deviation from the exact script is inevitable.

In the Cup Game, the performer had one page of script per trick, plus an extra page for the final questions, formatted as shown here.

(Spoken script in italic, directions non-italic in brackets.)
(First: check that they are okay to be video recorded.)
You can see three cups in front of you.
I've been training myself in order to hear the sounds created by the cups sliding - so I can tell where each cup moves without being able to see them.
So, put the object under the middle cup.
(wait for the object to be put there)
Okay, now, when you move the cups around, I want you to hold the cup like this, and move one cup at a time - this makes the noise recognisable to me.
Maybe once I've learnt a bit better I'll be able to do without that, but bear with me for now.
I'll put on this blindfold, and turn my back to you, so you can see I'm not cheating.
(turn round)
Okay, now move the cups a couple of times.
(wait for them to finish)
Lift up the cloth on my side, and pick the cup.
(hopefully I pick the right one here!)
ta- da.
That's the end of the first trick - so, what do you think happened there?
(need to get answers to questions)
...
Now, how I did that - Really, it was all a trick. I'll lift up the cloth, and you can see this computer, which is tracking the position of the cups, so it knows which one you put it under.

What is the Deception?



This is simply using a computer as a tool, to help in a deception played directly on the audience by the performer. The audience member doesn't know that a computer or camera are being used yet. The computer is hidden under the table, and the camera is not visible due to the blackcloth.

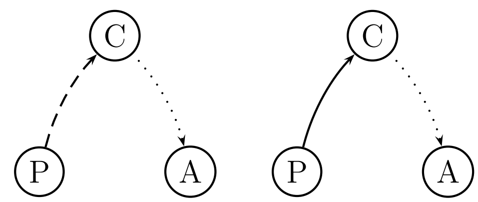
Desired Effect:
Audience thinks performer can tell which cup the ball is under in some magic way.

Questions:
Did the tracking work?
Did they spot the camera?
Do they believe the explanation?
Do they know there was a computer involved?

Exposure:
Black cloth raised to show computer / camera - and to set up for the next trick.

So, now you know about the tracking, I'll show you it in action.
(Takes the cup, and puts the ball under it)
Now, I put the ball under the cup, and press 'go'. When this happens, it tells me to shuffle the cups.
If you look at the screen now, as I move the cup with the ball under, the red circle on the screen moves.
(Moves the cup around - secretly swaps the ball out)
So, you can see as I move the cups around a bit, it keeps tracking.
Pay attention to the cups so you can keep track of which one the ball is under.
(Move the cups lots now - one with the ball in goes to the left hand side)
Which cup do you think the ball is under?
(Hopefully they point at the wrong one)
Now I'll let the computer have a go.
(Hopefully it picks the correct one!)
Lets see who won!
Ta da!
(pick up the cup to show it)
so, what do you think happened there?

What is the deception?



This involves two deceptions. Firstly, the computer knows that it is going to be fooled. The video image is faded out, so that when the person is looking at the cup tracking output, they can't see what the performer's hands are doing. Demonstrating the tracking is used as a misdirection for the audience (which the computer plays along with, as it knows the trick is occurring). The second deception is a simple hard-coding of which cup the computer will choose - just controlling the computer to make it fool the audience. The performer, who knows which cup the ball is under, can make sure that this cup appears in a particular position.

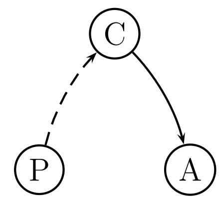
Desired Effect:
Audience thinks performer has moved the cups too quickly for them, but computer has tracked okay.

Questions:
Did the screen display work as a misdirection?
Did they pick the wrong cup?
Did the computer pick the right cup?
Do they think that there was a trick, or that the computer tracked the objects better than them.

Exposure:
This will be exposed as the final trick.

Now you have a chance to try and beat the system. I've been using the computer to trick you all this time, so it's only fair that you get a go. This time, I'm going to act stupid, and just trust the computer, whilst you try to fool it.
Pop the ball under the middle cup, and then see if you can shuffle the cups to beat the computer.
(let them shuffle the cups for a bit - until the computer is about to guess)
Okay, stop, and let it guess.
(if the computer gets it right, or they get told off)
You can have another go if you like
(if they get it right too easily)
You obviously found that pretty easy - how about trying to fool it in another way?
(When they are bored / or manage to fool it)
Okay, you managed to fool it.
(some questions about how they did it)

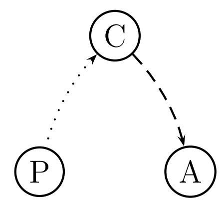
What is the Deception?



The computer knows that the participant is trying to fool it - so it tries its best to keep track and avoid being fooled.
Desired Effect:
The audience work out some way to fool the computer.
Questions:
Did they manage to fool the computer?
How many tries did it take?
What strategies did they use to fool the computer?
Did they exploit the weaknesses of the computer when trying to fool it, or do they fool it in the same way as they would a person?
Does having technical knowledge make it easier to fool?
Exposure:
Not really a trick

Now, I'm going to show you how I tricked you.
While you were looking at the screen, at the very start of the trick, I switched the ball. The computer can't track the ball, only the cups, so all you could see on the screen, is the ball itself. So, once it is switched, the tracking is tracking the wrong ball.
(Demonstrate the switching)
However when I did this, I let the computer know that I was fooling it, and told it to always guess the left hand cup at the end. This meant that I could make sure it gets the correct cup, rather than the one suggested by the tracking.
Whereas if I hadn't cheated, it would have kept tracking the wrong cup, and guessed wrong.
(Demonstrate this)
You can also fool it in some other ways - originally the ones I thought of were:-
1) by putting the ball under the wrong cup at the start.
2) by covering the cups from the top so the camera can't see them and switching.
3) by turning off the light, so it can't see at all.
4) by moving the cups out of view.
However, I designed the program to try and detect the last three - 'which is why it caught you'? .

What is the Deception?



The performer knows how to fool the computer, as they know the limitations of the tracking algorithms - in particular that it doesn't track the ball, just the cups.
Desired Effect:
The audience understands more about how this kind of system works.
Exposure:
This is itself the exposure of the whole of the rest of the tricks.

Okay, that's the end of the tricks - now if it's okay, I'd just like to ask you a few questions.

1) Did you enjoy the tricks? How did you feel about being tricked by the computer? Was it a disappointment to discover that there was a computer, or to discover how simple the second trick was? How about cheating the computer yourself -

few q's about you:

2) What do you do for a living?

3) How often do you use computers?

4) What would you consider your level of computer expertise to be

- have you ever programmed a computer?

5) This experiment used computer vision, where the computer uses a camera to see what is going on, in this case the movements of the cups. Were you aware of this kind of technology before today?

6) How often have you seen magic performed? Do you watch it regularly? Have you seen it live much? Have you ever performed a magic trick yourself?

7)One last thing, when I've finished these experiments, I will probably publish something about it, are you okay with me using clips of the video including you, or pictures from it, or quoting things you said? No problems if not.

(get this in writing, on the permissions form)

Appendix C

Pressing the Flesh: Sensing Multiple Touch and Finger Pressure on Arbitrary Surfaces

Joe Marshall, Tony Pridmore, Mike Pound, Steve Benford & Boriana Koleva

Presented at Pervasive 2008 (Sydney, Australia)

{jqm,tpp,mpp,sdb,bnk}@cs.nott.ac.uk

This paper identifies a new physical correlate of finger pressure that can be detected and measured visually in a wide variety of situations. When a human finger is pressed onto a hard object the flesh is compressed between two rigid surfaces: the surface of the target object and the fingernail. This forces blood out of the vessels in the fingertip, changing its colour slightly, but systematically. The effect is visible to the naked eye and can be measured using techniques from computer vision. As measurements are made of properties of the hand, and not the target surface, multiple-touch and pressure sensing can be added to a range of surfaces - including opaque, transparent, smooth, textured and non-planar examples - without modification of the underlying physical object. The proposed approach allows touch sensing to be fitted to surfaces unsuitable for previous technologies, and objects which cannot be altered, without forfeiting the extra range of expression of pressure sensitivity. The methods involved are simple to set up and low cost, requiring only a domestic-quality camera and a typical computer in order to augment a surface. Two systems which exploit this cue to generate a response to pressure are presented, along with a case study of an interactive art installation contructed using the resulting technology. Initial experiments are reported which suggest that visual monitoring of finger colour will support recogntion of push events.

C.1 Introduction

Touch sensitive surfaces, such as graphics tablets, interactive whiteboards, touch screens etc. have existed for some time. Touch sensitivity, however, typically requires the surface to be enhanced with some kind of embedded electronics, or in the case of capacitive sensing on glass [12], to have electronics below the surface. Computer vision has the potential to create touch interfaces without embedding electronics in the target surface, and also to detect multiple touches. Current systems, however, are typically either unable to detect the difference between touching and moving a hand or object near the surface [2], or can only detect the presence of a finger or object next to the surface (by using cameras at right angles to the surface [8], or multiple cameras and some form of 3d disparity measurement [6, 17]).

The main contribution of this paper is to identify a new physical correlate of finger pressure that can be detected and measured visually in a wide variety of situations. When a human finger is pressed onto a hard object the flesh is compressed between two rigid surfaces: the surface of the target object and the fingernail. This forces blood out of the vessels in the fingertip, changing its colour slightly, but systematically. Increased pressure increases the effect, up to a limit determined by the thickness of flesh on the finger. Colour change may be seen either by examining the pattern of colours in the fingernail or, if the target surface is transparent, by looking at the fingertip through the surface. When viewed through a transparent target surface, increasing pressure increases the amount of flesh from which blood is expelled, creating a larger region of paler skin. When viewed from above the hand, through the nail, increased pressure forces more blood to the base of the nail, concentrating colour there. Both these events are clearly visible to the naked eye.

In what follows we describe computer vision-based sensing methods which exploit this cue. As the approach relies on measurements of the physical properties of the hand, and not the target surface, it has the potential to add multiple-touch and pressure sensing to a range of surfaces - including opaque, transparent, smooth, textured and non-planar examples - without modification of the underlying physical object. This allows for many new items, such as stone carvings or wood, to become touch sensitive interfaces. No technology need be embedded into the target object; all that is required is that a colour camera be positioned to view the effect. The proposed method is therefore potentially highly flexible, easy to install, low cost and portable. It requires only a domestic-quality camera and a typical computer in order to augment a surface.

The proposed approach is expected to be of particular use in environments such as museums, science centres and galleries. Here, visual sensing of finger colour can allow people to interact directly with existing physical objects, or with glass cabinets containing objects of interest, without having to customise the objects or cabinets themselves. The method has benefits for installation designers, allowing museum and science centre staff to construct interactive exhibits and environments based upon their existing catalogue of objects. For example, historic tools in a countryside museum could be touched in order to trigger audiovisual material about their use. The flexibility of the approach means that exhibits could also be reconfigured easily and on a regular basis, maintaining visitor interest.

The paper is organised as follows. Section 2 briefly reviews relevant prior work, before Section 3 describes finger pressure sensing by viewing the tip of the fingernail from above. Section 4 then describes a method which uses the same visual cue, but views the hand from the rear of a transparent glass surface. The fingernail-based method was used to create an interactive art installation, in which the user was able to interact with a pressure and touch sensitive rock. This installation is presented as a case study in Section 5. A key motivation for the development of touch sensitive interfaces is the ability to detect touch events such as contact, pushes, taps, etc. Initial experiments are reported which suggest that visual monitoring of finger colour will support recogntion of push events are described in Section 6. Finally, conclusions are drawn in Section 7.

C.2 Prior Work

Touch sensitive graphics tablets [15] and touch screens [5] are widely available pressure sensing interfaces. The most common pressure sensitive interface in production is the laptop touch pad. Whilst these are typically used purely for on/off touch pressure, most, such as those made by Synaptics [13] also are able to detect variations in pressure. Some models, for example the Mitsubishi DiamondTouch table [3], even allow the detection of multiple touches, although not pressure.

Following a different approach, Schmidt et al. [9] used load sensors to create a touch sensitive table. In addition to touch events, they detect several contextual events such as objects being put down on a table, which is interesting as it relates to our goal of augmenting existing objects. Schmidt et al, however, cannot support interactions with objects other than moving them around on the table. This does not require additional technology in the sensed object itself, but load cells are required to be fitted at four corners of the surface. It is also limited to single touch interaction on horizontal surfaces.

Exploiting vision and related technology to create a touch detecting screen is not a new idea. Various methods have been used, such as scanning laser rangefinders [10], internal reflection inside a glass plane[4], multiple cameras and planar homographies to detect only pixels that are near the screen [17], and the visual detection of (somewhat exaggerated) finger gestures in order to detect touches on a virtual keypad [14]. These visual methods typically fail to detect pressure differences during touching, although some level of pressure sensing has been demonstrated with the internal reflection method, by using the size of the finger's contact area. The finger surface area is also used in Benko et al's multi touch table[1].

Figure C.1: Nail at different pressures

Benko et al suggest it is too innacurate to detect pushing reliably and define a special rocking gesture for clicking which their system is able to detect. The two sided LucidTouch system[16] also uses visual tracking to detect the hand position, however it uses a separate touch sensitive pad in order to to detect touches on the surface (as the vision tracking method used is unable to detect touch).

These technologies are designed for use in interactive whiteboard, wall display or table interfaces. They usually require modification of the sensing surfaces in some way, or place restrictions on the surface being monitored. They are also currently designed for completely flat user interfaces. This may be suitable when used as an interface to standard GUI style applications; however as interface designers move beyond the GUI, this may become a limitation. When augmenting existing objects, it is hard to guarantee complete flatness. Bumpy or angled surfaces may also be useful to allow tactile feedback as to where the hands are, which is commonly seen as a reason why touchscreen interfaces such as virtual keyboards have only had success in niche applications.

C.3 Fingernail Sensing

When the fingertip is pressed down on a surface, the blood under the nail concentrates at the bottom of the nail, and the tip of the nail becomes whiter (Fig. C.1). This effect is very consistent, and only requires a small amount of finger pressure for a difference to be clearly visible to a human observer. This section discusses the automatic visual detection of this cue.

The fingernail sensing system uses a basic background segmentation algorithm, followed by a contour detection operation to find the fingertips. When a finger is detected which has not moved more than a small threshold since the last frame, the image of the fingertip is examined, and the distribution of colour in the nail quantified. This reflects the pressure exerted by that finger.

Initial attempts at sensing pressure used the two parts of the nail, the tip and the bottom, and compared the colours of these to detect a change. However, the exact location of the white areas on the tip of the finger proved to vary significantly between individuals, and is also difficult to sense from any distance. For example, when viewed from 60cm with a 320x240 pixel camera, the nail is approximately

10x12 pixels in size, which means that the tip area in particular is too small. However, while the fingernail is almost uniformly coloured when no pressure is applied, two distinct colours appear on the nail when pressure is exerted. Because of this, rather than use located features on the nail, we simply take the variance of the hue of the pixels in the nail area. In order to calculate a mean hue, the hue is represented as a 2d vector, and an arctangent applied to this. Variance is calculated with an allowance for the circular nature of the hue metric.

$$MeanHue = \text{atan2}([\sum_{1}^{n}\cos(Hue)], [\sum_{1}^{n}\sin(Hue)])$$

$$VarHue = \frac{1}{n}\sum_{1}^{n} min((Hue - MeanHue)^2, (360 - (Hue - MeanHue))^2)$$

Initial testing has shown this metric to relate strongly to finger pressure. It is also much more detectable at a distance, and produces similar results on different fingers. Variance of the brightness of the pixels can also produce useful data in some conditions, however it is, as might be expected, extremely sensitive to illumination changes. When the finger is pressed down hue variance clearly increases, with the opposite effect visible on release. Pushing less hard produces an intermediate response. Because the blood under the skin moves back into its normal place relatively slowly, there is a natural smoothing on the release of approx 100ms, this may be useful for 'debouncing' purposes, avoiding multiple presses being detected when the finger is only pushed down once.

It is also clear that depending on lighting and individual variation, the absolute variance values alter somewhat. A floating normalisation window is therefore employed, with the value of 'pressure' detected being mapped to 0...256, by using previously recorded pressure values as a max and min. A constant minimum pressure range (mr) is used, for the case when the finger is first seen, and only a small amount of data is in the window. This avoids large random fluctuations if the finger is simply placed down and not pressed. The raw to normalised conversion is expressed as:

$$normalised_t = \frac{256 * (raw_t - \left(\min_{k=t-windowSize}^{k=t}(raw_k)\right)}{\max\left[mr, \left(\max_{k=t-windowSize}^{k=t}(raw_k) - \min_{k=t-windowSize}^{k=t-1}(raw_k)\right)\right]}$$

This assumes that when the finger is first seen, there is no pressure on it, which is the case in typical use; even if a press is occurring the finger is first seen as the press starts. This conversion, whilst it means that no exact pressure data is available, makes push and release events clearly visible, and allows intermediate pressure values to be acquired. Normalisation is effective as long as the raw variance is altering with finger pressure. It has a compressing effect on the raw curves, which was desirable in our application (Section 5), but may or may not be suitable depending on context.

C.3.1 Initial Evaluation

In order to test the fingernail algorithm, a test rig was constructed, with the user's finger pushing on an electronic scale which served as a ground truth pressure gauge. The output from the scale was then video recorded along with the output from several brief sessions of pushing and releasing a single finger. The test rig was able to detect a 'weight' of 2kg (a force of approximately 19.6 Newton). In practice, this limit was not a problem, as forces outside this range proved uncomfortable to apply. The scale reported weight with a relatively slow update rate, updating at up to 4 times a second. Measurements from the visual system were taken each time the scale's reported weight changed. The system was run, and the hand moved into view until the hand tracking found the finger, and then the output was recorded for approx 50 seconds.

C.3.1.1 Single User Reliability

Once the data had been recorded, the raw hue variance data for each test session was scaled in order to make the mean and standard deviations the same as the ground truth. These normalised graphs showed a very good fit to the ground truth data, with a certain amount of clipping at the highest pressures in some tests. These results were analysed using regression analysis, which gave a P value of $<0.1\%$ for all subjects. Figure C.2 shows 3 different user's normalised pressure outputs plotted against ground truth. These graphs demonstrate the ability of the algorithm to reflect several push release cycles accurately. Detecting an initial push is possible, as the minimum pressure range means that an initial push will have a different profile to just touching (the first graph in figure 3 starts by just touching the surface, whereas the second and third graphs start with a push, the visual measure ramps up high straight away. The tracking works at 30 frames per second, limited by the camera frame rate, rather than any processing constraint, which is fast enough to detect quick push and release cycles.

C.3.1.2 Between User Variation and Lighting Variation

Tests were carried out on different days, in a naturally lit room. This meant that the system was exposed to some lighting variation. To quantify the possible effects of lighting variation, one user was tested on two different days, both times using the same finger.

The variation in lighting had a major effect on the raw variance values from the system, with the same user showing a significantly lower range of variances, which were also significantly higher than their previously recorded values. Multiple users

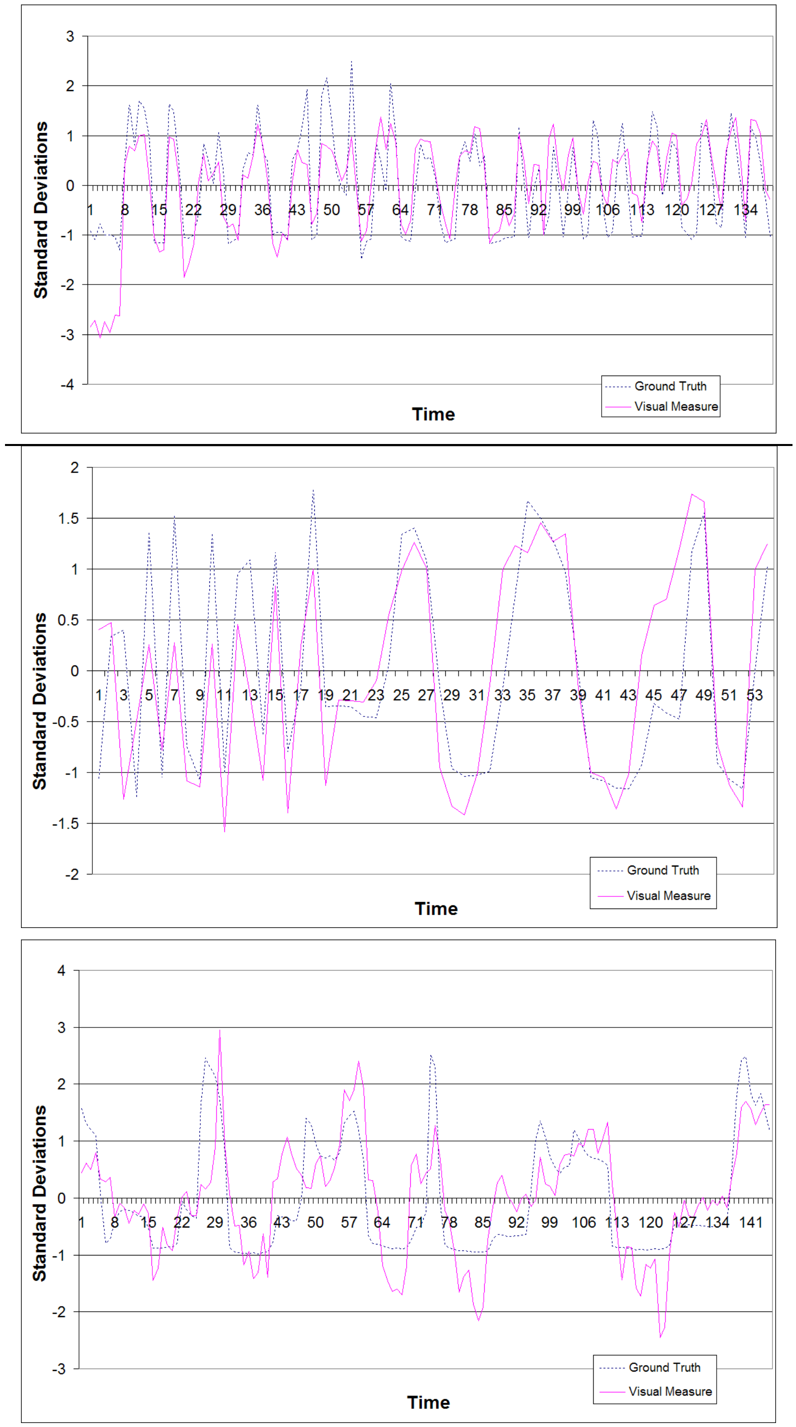


Figure C.2: Examples of Performance of Pressure Tracker over Time

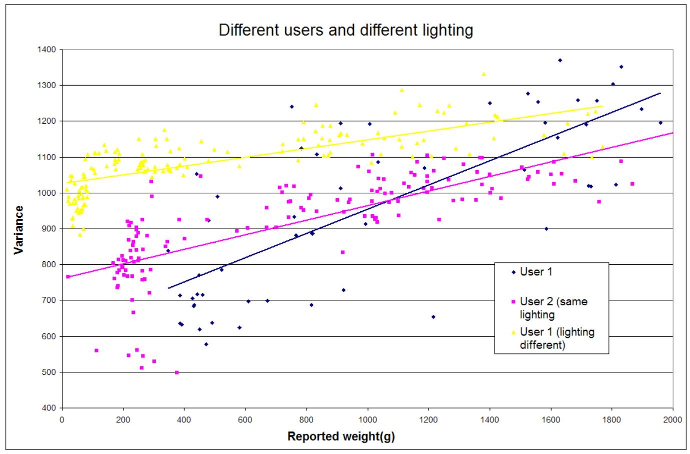


Figure C.3: The Effect of Users and Lighting on Fingertip Variance

in the same lighting conditions also had differences in the distributions of variance, although these were significantly less than the lighting induced variances. Figure C.3 shows some examples of these effects.

These two factors mean that unless very controlled lighting is available, and a training session is undergone for each user, this method is not suitable for providing absolute pressure information, ie. it is not suitable to replace a load sensor. However, when normalised as described above, it can be employed in interfaces where a correlate of pressure, rather than true pressure, is required. It seems likely that colour-based measures can support detection of more fuzzy actions such as pressing, pressing hard, pressing softly etc., as is required in most touch based interfaces.

C.3.1.3 When Does This Work?

Several factors may cause this method to fail. Firstly, nail varnish or gloves will obviously cause the system to fail, as the fingernail cannot be seen. Secondly, if the fingernail is very brightly lit by direct sunlight, this may reflect off the nail, making it impossible to see the skin colour beneath it. This was the case during one of the test sessions, with the system failing to work until a curtain was drawn to block the bright rays of sunlight.

The method is reasonably robust to changes in finger orientation. As long as the length of the fingertip can be seen, a correlate of pressure is produced and changes in the hue varaince reflect changes in pressure. If the angle is changed during sensing however, the values can be seen to change slightly. This means that there is potential for use on non-flat surfaces, as long as the finger is not changing in angle massively during a single touch movement. A slight side effect of our simple hand tracking system is that when the fingers are clasped round an object, so the fingernails are out of view, the knuckles and what is visible of the finger above them are detected to be fingertips by the system. When the knuckles are detected as fingertips, the system still responds, as the knuckles are differently coloured to the rest of the finger, and grasping causes the ratio between the knuckles and the part of the finger that is visible to change, thus altering the variance of the detected ‘fingertip’ (see Figure C.4) . Potentially useful data is also provided if the hand is held in the air, and the thumb is squeezed against the bottom of a finger.

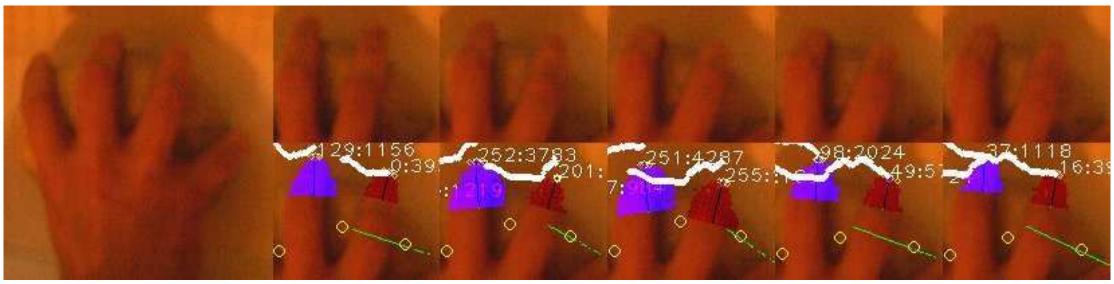


Figure C.4: Grasping and releasing an object. Each frame shows the zoomed in middle finger, and the pressure graphs next to the middle and index fingers

C.4 Skin on Glass

To assess the potential of visual monitoring of skin colour to detect pressure on transparent surfaces, such as windows, glass cabinets, etc., the same approach was applied from the other direction, tracking the finger through a sheet of glass. Changes in skin colour were recorded as the finger was pressed against the glass. It was found that at the point of contact pressure was sensed reliably if normalised as described in Section 3. Intermediate pressures were again detectable, and pushing and not pushing generated distinct output profiles. Sensing through glass may be of particular value as it allows the computer and camera to be entirely enclosed, for example behind a shop window, or inside a glass case, with no exposed electronic parts. It also has an advantage over the fingernail tracking in that it is less susceptible to occlusion, which may be a problem in some uses of the fingernail method.

This technique is, however, not quite as reliable as the fingernail-based method in one particular: until the hand is touching the glass, sensing is somewhat erratic. Further research is required, but this is probably the result of the changing distance between the glass surface and user’s hand. It seems likely that the fingernail-based method is more reliable because the nail is firmly attached to the surface of the finger, so that the relationship between the fingertip and the surface through which it is viewed remains constant. In the test application, the effect is reduced by only starting to record pressure once the detected fingertip has been in the same position for 3 frames. This means that there is a delay of approx 1/10th of a second before continuous pressure readings begin. It also means that if the hand is held very still in the air in front of an interface, pressure sensing will begin, although it will only break if the hand is very slowly moved directly towards the camera, which proved hard to do in testing. An output from this sensing during two pushes on a rather dirty and reflective sheet of glass is shown in Fig. C.5 (this version was tested with a black background, as skin segmentation did not prove a problem in the initial skin on glass tests). The method works well even in sub-optimal conditions. Figure C.5 also shows a graph showing the comparison of the skin on glass to the ground truth measurement (as used in Figure C.2 for evaluation of the fingernail tracker.) Note that this system uses an identical algorithm to the fingernail tracking, with the normalisation taking care of the smaller absolute variance values seen in this method. It is possible for a user to simply move their hand to the other side of the glass and use the fingernail tracking without any recalibration.

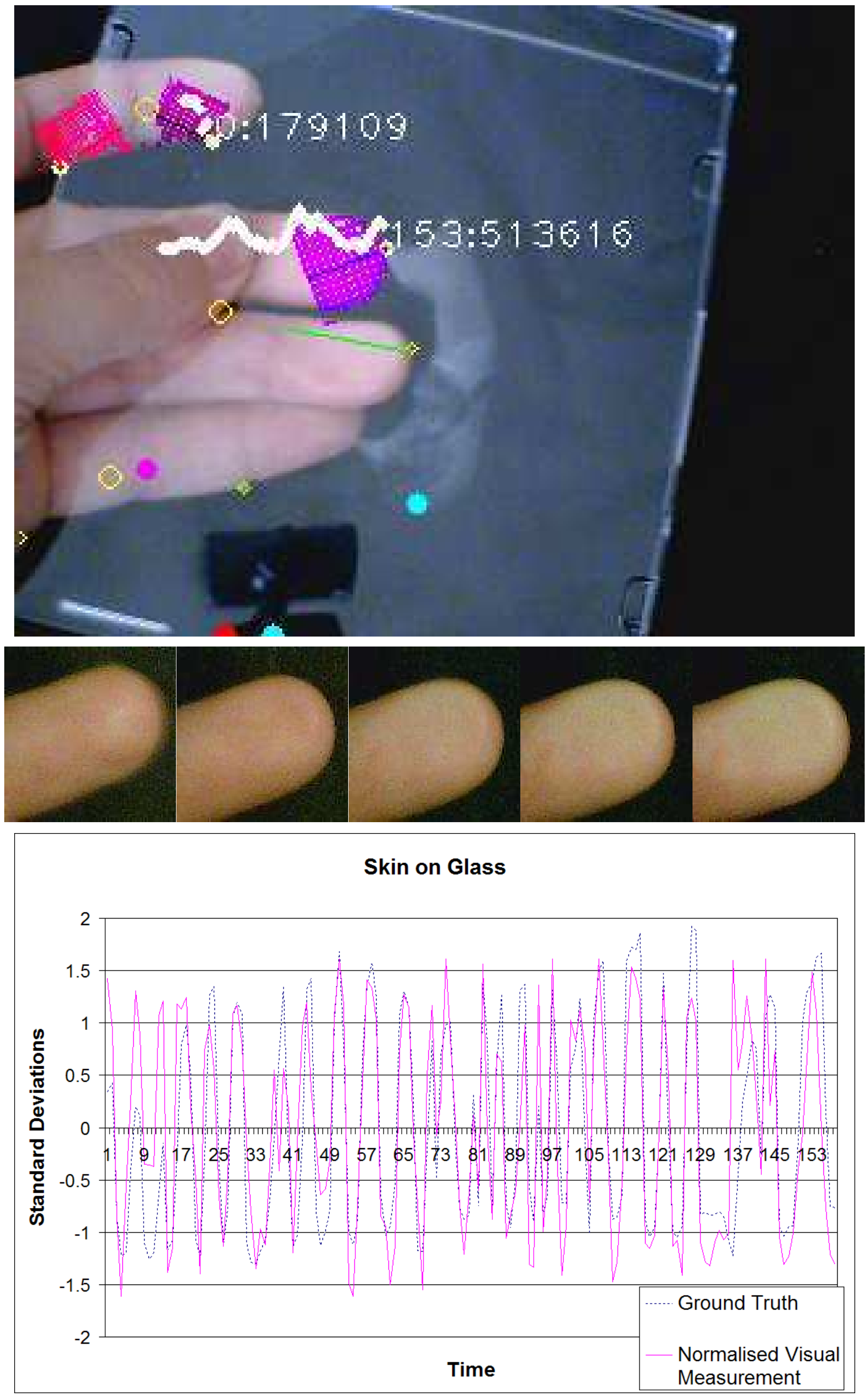


Figure C.5: Sensing pressure through glass - the line of pictures shows the fingertip over a single push sequence.

C.5 Case Study: Rock

The fingernail tracking algorithm was evaluated further in an interactive art installation called Rock. The installation presents a rock in a cage as a pet. A web camera is attached to the cage's top, and a computer and speakers are hidden under a table that the cage is on (see Fig. C.6). Rock uses gestural input and audio output to mimic the personality of a small pet rodent such as a guinea pig. It is designed to have quite a timid personality and to be easily frightened. The rock is an extreme test of the behaviour of the fingernail algorithm with a large range of gestures and angles, and provides a testing ground for graceful fallback in situations where it is impossible to sense pressure.

Initially the rock makes a quiet steady heartbeat sound. When the rock is touched it responds by making animal sounds, and the heartbeat changes to signify its level of fear. Touching the rock in different ways can provoke varying responses, for example if touched gently and slowly, it is likely to make quiet purring noises and not be very scared: grabbing at the rock too quickly scares it and makes it snarl or growl.

The rock is designed as an ambient installation, to be left in a gallery or space at an event, and interacted with by people with a minimum of direction. As such, it is designed to attract people to interact with it; this takes two forms. Firstly the heartbeat sound attracts interest to the rock when it is not being interacted with. Secondly, the interaction with the rock is designed to be interesting to onlookers. The interaction is designed so that onlookers can see part of the way that the rock is being interacted with, but so that part of the interaction is not visible to them. In particular, the finger pressure detection is used here and provides an aspect of the interaction that is unclear to onlookers, and designed to intrigue people into interacting with the rock themselves.

Figure C.6: The Physical Setup of the Rock

C.5.1 Technology implementation

The single camera on the top of the cage is the only input mechanism for the rock. The camera is carefully positioned so that it can see all of the bottom of the cage. All the output comes from the speakers, which are positioned so that the sound seems to come from the bottom of the cage.

The computer detects the silhouette of a hand reaching into the cage by use of a simple colour threshold to select pixels which match the colour of the background or the rock. There is deliberately no skin colour detection, or scene based background subtraction, in order to make the system responsive to non-skin objects put into the cage (as long as they are not the same colour as the background or the rock), and also to allow the rock to be moved in the cage without breaking the background model. This makes for a very reliable and simple detection of the hand silhouette when the hand is inside the cage.

The system detects how close the hand is to the rock, and uses this over time to calculate a measure of how fast the person's hand is approaching the rock when they reach into the cage. It also attempts to find the fingertips and detect the average finger pressure on the rock over all the fingertips it can see, by using the algorithm described in Section 3.

These two measures, of approach speed, and pressure are mapped respectively into two variables, 'fear', and 'excitement'. These variables are mapped onto a set of audio samples, and audio processing filters which alter these sounds. The audio samples used were made by one of the authors, and are categorised as to how scared and how excited they sound. The audio processing effects are a mixture of time and pitch shifting, and are used in order to make the sounds sound different every time they are played rather than like a fixed set of samples. Examples of sounds that the rock may make are a low growl if it is scared but not very excited, a high pitched snarl if it is scared and excited, purring sounds if it is not afraid but not very excited and squealing sounds if it is excited and not afraid. A slight element of randomness is added into this mapping; this is designed to make the rock be mostly predictable, but to avoid letting the users be certain how it will respond to a particular gesture. The heartbeat sound continues all the time beneath the animal sounds, getting faster and louder when the rock is more scared.

C.5.2 Testing the Rock

The rock was exhibited at a recent digital art conference. At this event, it was placed in a corner of a corridor space, where a lot of people were passing by, as an ambient installation during the conference. This allowed us to see the rock interact with approximately 100 people, from various backgrounds including art, architecture, sound design, HCI etc. The rock was running for over 9 hours, and was very successful in this environment; the installation was awarded best paper prize.

Initially, one of the authors was with the rock, introducing it as his pet, in order to entice people to play with it. After a few people had played with the rock, this became unnecessary, as people started bringing back other people to show it. At this point, the rock became more interesting, as the explanations people were creating for its behaviour became increasingly complex and rich. At the end of the event, one of the participants was very attached to the rock and even asked if she could take the rock home. The descriptions of the rock and its personality were very varied, ranging from 'cute' to 'strange' and 'disturbing'.

There were several ways in which people interacted with the rock. Most common initial interactions were poking it, either suddenly, or gingerly reaching in to touch it. In these cases, the technology responded reliably. Once people had realised that the rock was not going to bite them, they explored more complex interactions, such as stroking it (which worked as long as they didn't move their hands too fast), and grasping it. Grasping was interesting, because the effect discussed in Section C.3.1.3 meant that the knuckles were tracked, giving a pressure signal as to how hard the rock was grasped. This meant that in this (relatively common) mode of interaction with the rock, the tracking still worked, although slightly less reliably. A few people did things such as picking up the rock from underneath, waving their hands right in front of the camera, or closing their hand into a fist when touching the rock. In these situations, the pressure tracking broke, and the rock responded to the movements using only the silhouette of the visible part of the arm, which led to slightly unpredictable responses to these particular movements; it was important in the design of the rock's 'personality' that it handled the cases when finger tracking data became unavailable, and still provided some kind of response. The unreliability when presented with these odd gestures was translated in the user's eyes to become a facet of the Rock's personality, for example as it not liking having strange things done to it.

While observing the rock, it was clear that the balance between the unpredictable nature of its response to odd actions, and the predictable response to actions such as stroking gently and holding it, formed a part of the success of the installation. The ambiguity allowed people to spot 'patterns' and create expla-

Figure C.7: Interacting with the Rock

nations, and meant that whilst people could to some extent learn things about how to control the rock, such as not to grab at it and scare it, or by using gentle touches to make it happy, they were not able to get to a level where they felt they had complete control. One important thing however is that the level of reliability was such that the 'owner' of the rock was able to demonstrate that the rock 'liked' him, and that people were able to learn how to touch it to make it likely to make 'happy' sounds.

The finger pressure sensing method was important in this installation, as it allowed a very expressive mode of interaction with the rock, but without having to augment the rock with sensors. Within the constraints of the cage, this created effectively a wireless, remotely powered, touch sensitive moveable user interface, which was made of seamless stone, with no charging connectors or battery compartments. Alternative ways to create similar effects would have created points at which the audience's suspension of disbelief was broken. For example, a pressure sensor under the rock or cage would fail to work if the rock was lifted, adding sensors to the rock itself would be hard to do without external electronics, battery compartments etc. which would break the concept of it being an organic creature.

C.6 Using a Bayesian Classifier to Detect Push Events

A key motivation for the development of touch sensitive interfaces is the ability to detect touch events such as contact, pushes, taps and double clicks. To provide an initial indication of the feasibility of detecting such events given colour variance data a bayesian classifier was implemented and used to detect contact between hand and surface. Colour-based detection and location of human skin is now commonplace in computer vision systems. A number of skin detection techniques have been reported [7, 18, 19, 11], most based upon the work of McKenna et al. [7] which showed that colour spaces exist in which, for a wide range of nationalities and ethnic backgrounds, human skin is tightly clustered. The classifier to detect contact between human fingers and a target surface by identifying compressed flesh adopted a similar approach.

A camera was placed behind a sheet of non-reflective, smear-resistant glass, providing a clear view of the user's hand as s/he interacted with the other side of the surface. The challenge was to use the resulting colour images to recognise the differences between:

- the normal, i.e. uncompressed, skin seen when the user's hand is in view, but not in contact with the glass;
- the compressed flesh that appears when the user touches the glass surface;
- the environment behind the user.

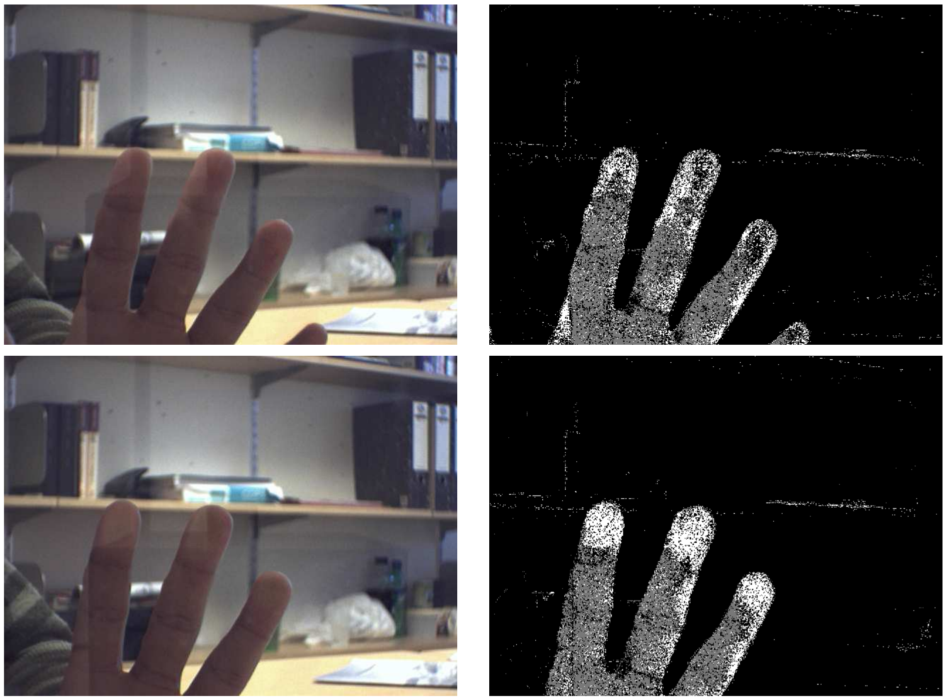


Figure C.8: Skin on glass - hand just touching glass (1), and fingers pushed against (2) - grey pixels are classified as non compressed skin, white pixels are compressed skin.

All experiments were carried out in an office/laboratory environment, so the background comprised arbitrary coloured objects. Some of these objects were approximately skin-coloured, but no other people (i.e. no additional real skin) was allowed into the field of view. Six individuals, of mixed age, sex and race took part. Each was first asked to press his/her hand flat onto the glass panel to provide easily identifiable examples of contact, and then invited to press, tap, or otherwise touch the glass at will. Two minutes video of each subject was captured and analysed off-line.

Following [7, 19], the well-known hue-saturation-intensity (HSI) colour space was employed throughout. The hue (H) and saturation (S) values associated with human skin are known to cluster tightly, though intensity (I) varies widely. Bayesian classification was used to separate the three classes (uncompressed skin, compressed skin, non-skin) identified above as shown in Fig C.8. Models, in the form of approximations to probability density functions for each class were first constructed from manually identified training data. This classification of pixels into tip and non-tip could potentially allow for reliable detection of touch pressure, by detecting the size of the compressed region of the fingertip. Raw colour values were examined to determine whether or not the information required was present, before any features or summary statistics were computed, in the base image data.

When trained on data from an individual's hand, bayesian classification was

Strengths	**Weaknesses**
No modification of tracking surface required	Viewpoint and occlusion
Quick, easy and cheap setup	Lighting
Smooth pressure sensing	Relies on hand tracking
Potential to support automatic detection of touch events	Not fully 3D
Wide range of surfaces can be augmented	
Multiple touch	

Table C.1: Benefits and Challenges of using Fingertip sensing

found to be effective. Contact between fingertip and glass could be reliably detected. However when applying the same algorithm to multiple users, by pooling training data to produce composite colour models, it was found that the variation between users was often equal to the difference between compressed and non-compressed skin for a single user. The experiment therefore demonstrated that variations in individuals' skin colour can better support automatic detection of touch events. As a result, further work on event classifiers will exploit time-based measurements of individuals' finger pressure, similar to those we have used for smooth pressure sensing.

C.7 Conclusions

We have demonstrated a new correlate of finger pressure which can be measured visually using standard equipment in a wide variety of circumstances. The method detects compression of the fingertip by monitoring changes in the colour of either the skin or the fingernail. Table C.1 summarises the main strengths and weaknesses of this approach to pressure sensing.

The method allows the addition of an extra dimension of expressiveness to previous vision based hand & finger sensing systems, without requiring complex addons such as multiple cameras, or augmenting the surface in any way. It is inherently multiple touch, as it measures a feature of the pressure on the finger, rather than the pressure on the surface below the finger. It is quick, easy and cheap to setup. The technique extends the range of materials and surfaces available to standard pressure sensing, by allowing touch pads to be created from any relatively firm surface which is visible to a camera. As demonstrated in the Rock example, fingertip pressure sensing does not even require a flat surface, working well when given a bumpy surface. Though further development is required to produce a working system, the colour measures employed here clearly have the potential to support detection of a variety of touch events.

The approach is, like many vision-based techniques, potentially sensitive to camera viewpoint and occlusion and is unlikely to work well in some extreme lighting conditions (very bright sunlight & darkness). It is also reliant on the hand tracking working correctly in order to function; if the tracking fails, no pressure sensing can occur. It is not fully 3D, as it cannot sense pressure when the whole of the fingers are out of view, so applications have to be designed to degrade gracefully if this is a possibility, however it provides useful data in a large range of situations, such as when the fingertips themselves are out of view, as long as the grasping hand and the rest of the fingers are still visible to the system. Also, whilst it works on a wider range of surfaces than most current systems, there clearly is a limit to what surfaces it can work reliably on, for example surfaces such as cushions, gels or liquids will all be impossible to augment.

C.7.1 Potential Applications

Whilst this technology clearly may be useful in tabletop displays and other common multi-touch interfaces, it has most to offer in the creative, museum and educational sectors. The ability to augment an existing, everyday, physical object would be of particular use to museum, science centres, exploratoria and other similar places where a hands-on approach is encouraged. The skin on glass method of touch sensing provides a useful extra mechanism for objects which are in cabinets and unable to be directly touched. In this situation the hardware would be fully enclosed within the cabinet, which may be an advantage. Augmenting unexpected surfaces in this way has proven interesting and surprising to users in our case study; it is envisaged that in a museum setting, being able to augment the object rather than having a separate interactive display may provide a more direct and engaging experience.

As well as being useful for currently impractical applications, the techniques reported here make pressure and touch sensing available with a significantly lower setup time than existing methods and require no custom equipment; the Rock takes approximately 5 minutes to install and uses a cheap domestic webcam and PC. This means that the proposed method has the potential to be incorporated in mass market entertainment software, for example this could enable innovative interfaces such as used on the Nintendo DS touch screen game console to be created on a larger scale for home users (For example in Warioware Touched, users have to ‘rub out’ on-screen pictures, stroke dogs, whack moles etc. by using touch gestures).

C.7.2 Future Work

The work described here has demonstrated the potential of visual monitoring of skin colour to reflect finger pressure in a range of situations. Topics for future research include:

- investigation of alternative methods of capturing changes in skin colour, and their relation to finger pressure. In particular, though the current method is reasonably robust to changes in finger orientation it is not invariant under such changes.

- evaluation of the usability of the approach in a wider variety of application domains and scenarios, focusing on the creative, museum and educational sectors
- techniques for the automatic recognition of single and multiple touch events and gestures. As well as the gestures commonly used in GUI applications such as clicking and dragging, the work with the Rock demonstrated the possibility of detecting more unusual gestures such as grasping and stroking, which may be of interest for those designing applications which do not fit a standard desktop paradigm.With the addition of a more sophisticated hand tracker, it may be possible to further improve the tracking, by tracking touch actions using hand shape as well as fingertip cues, although it is not currently clear whether these may require per-individual training.

Acknowledgements This work is funded by the EPSRC.

Bibliography

[1] Benko, H., Wilson, A. D., and Baudisch, P. Precise Selection Techniques for Multi-Touch Screens. In *CHI*, 2006.

[2] Bérard, F. The Magic Table: Computer-Vision Based Augmentation of a Whiteboard for Creative Meetings. In *IEEE workshop on Projector-Camera Systems*, 2003.

[3] Dietz, P. and Leigh, D. DiamondTouch: A Multi-User Touch Technology. In *UIST*, 2003.

[4] Han, J. Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST*, 2005.

[5] MagicTouch. *Touch Screens*. `http://www.magictouch.com`.

[6] Malik, S. and Laszlo, J. Visual Touchpad: A Two-handed Gestural Input Device. In *ACM Int. Conference on Multimodel Interfaces*, 2004.

[7] McKenna, S., Gong, S., and Raja, Y. Face Recognition in Dynamic Scenes. In *British Machine Vision Conference*, 1997.

[8] Morrison, G. D. A CMOS camera-based man-machine input device for large-format interactive displays. In *SIGGRAPH*, 2007.

[9] Schmidt, A., Strohbach, M., van Laerhoven, K., Friday, A., and Gellersen, H.-W. Context Acquisition Based on Load Sensing. In *UbiComp*, 2002.

[10] Strickon, J. and Paradiso, J. Tracking Hands Above Large Interactive Surfaces with a Low-Cost Scanning Laser Rangefinder. In *CHI*, 1998.

[11] Störring, M., Andersen, H. J., and Granum, E. Skin colour detection under changing lighting conditions. In *7th Symposium on Intelligent Robotics Systems*, 1999.

[12] Synaptics. *Capacitive Sensing Technical Description*. `http://www.synaptics.com/technology/tcps.cfm`.

[13] Synaptics. *Touch Pads*. `http://www.synaptics.com`.

[14] Tosas, M. and Li, B. Virtual Touch Screen for Mixed Reality. In *Proc. of ECCV Workshop on HCI,*, 2004.

[15] Wacom. *Graphics Tablets*. `http://www.wacom.com`.

[16] Wigdor, D., Forlines, C., Baudisch, P., Barnwell, J., and Shen, C. LucidTouch: A See-Through Mobile Device. In *UIST*, 2007.

[17] Wilson, A. TouchLight: An Imaging Touch Screen and Display for Gesture-Based Interaction. In *Int. Conf. on Multimodal Interfaces*, 2004.

[18] Yang, J., Fu, Z., Tan, T., and Hu, W. Skin Color Detection Using Multiple Cues. In *Int. Conference on Pattern Recognition*, 2004.

[19] Zarit, B. D., Super, B. J., and Quek, F. K. H. Comparison of five colour models in skin pixel classification. In *ICCC International Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems*, 1999.