Avatar Mirrors: An inquiry into presence in facial motion capture animated avatars.

Donald A. Smith

An exegesis submitted to

Auckland University of Technology

In partial fulfilment of the requirements for the degree of

Master of Creative Technology

2015

School of Education

Table of Contents

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgments), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

List of Figures

Acknowledgements

I would first like to thank my research colleagues and staff at Colab AUT for support, encouragement, and always presenting new viewpoints on sticky challenges.

I would also like to thank my parents for their support of this project and faith in my tenacity, and my partner ZW for always giving support when most needed.

And a big thanks to my supervisor James Charlton for always going the extra mile; challenging, questioning, and making it happen.

Abstract

Avatar Mirrors: An inquiry into presence in facial motion capture animated avatars.

This study explores factors responsible for a higher sense of *presence* in a Computer Generated (CG) avatar animated using real time facial motion capture technology. The intent of this research is to maximise presence in the user and to minimise feelings of uncanniness.

Facial motion capture and CG animated faces are used to explore the relationship between participants, their embodiment in a facially imitative avatar and the virtual environment that it inhabits. The research proposes new perspectives on how avatars might be embodied by participants and offers new viewpoints and types of *mirrors* for reflection. The current "selfie" fad enabled by the latest mobile phone technology is treated as the equivalent to the myth of Narcissus in which the young hunter becomes entranced with his own image in a pool of water. Both are seen as manifestations of our fascination with our own image and indicative of a desire to better see and understand our own image.

The research found that increased fidelity and responsiveness of sensor coupling; in this case real time facial motion capture, along with manipulations of the virtual environment caused participants to report feelings of presence. Social presence; the feelings of presence of another intelligence, was explored with the use of high fidelity data scans producing convincing morphology and motion detection of user's contributing to engagement.

Introduction

This research is an exploration of facially imitative CG avatars and how we engage with them. Although other research has been done on the social aspects of CG avatars in on-line chat spaces and video games (Cleland, 2008; Tinwell, Grimshaw, Nabi, & Williams, 2011) there were no studies found documenting modes of interaction with a facially imitative motion captured avatar.

Avatars and our obsession with our facial image permeate many aspects of contemporary culture. From the selfie and banning of the selfie stick in museums (Gilbert, J, n.d.), to the popular "Avatar" by James Cameron ("James Cameron's Avatar," n.d.), we seem to be fixeated on actively engaging with technology in an attempt to represent ourselves. This study sees parallels between the evolution of CG avatars and our obsession with selfie enabling technology and self representation.

Although Steuer's "Defining Virtual Reality: Dimensions Determining Telepresence" (1992) was theorized more than twenty years ago; technological progress has been slow to deliver the mediated experience Steuer argues is possible. In his journal article "The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments" Biocca defines the term presence in a virtual world as referring to "the illusion of 'being there' whether or not 'there' exists in physical space or not" (Biocca, 1997, p. 13). As it can be argued that presence is a multi-faceted phenomenon (Kim & Biocca, 1997), this study will discuss terms of presence in context of definitions from works published by Biocca (1997), and "Embodiment and presence in virtual worlds: a review" published by Schultze (2010), to reference participant's embodiment in the virtual world.

Advances in technology that have enabled cinematic producers to create believable environments and characters using CG imagery come at the expense of huge computing power and are not achieved in real time. Real time applications such as video games are achieving advances in this area, but an autonomous avatar experience is still constrained by computing power and lack of Artificial Intelligence application software.

The road to believable CG characters is paved with notable failures where cinematic features, the leading edge of development, have lacked affinity and have become uncanny to viewers (Perry, 2014). Our ability to communicate utilising embodiment in a CG avatar will be limited or enhanced by our evaluation of the humanness of the character.

 Biocca comments that one of three most pressing design issues with regards to designing virtual humans or avatars is "… the engineering of an expressive face from the 3D geometry of avatars and agents" (Biocca, 1997, p. 10). CG avatars have typically used low numbers of polygons and simple shading models to compensate for internet bandwidth limitations and low computational power of consumer devices at the expense of the humaness of the CG avatar. Computing power may overcome impediments to a believable real time avatar but study is needed to identify factors affecting our perception of CG rendered characters with synthesised facial expressions.

This study is made possible by an emergent technology; real time high fidelity marker-less facial motion capture by Faceshift (Li, Weise, & Pauly, 2010; Weise, Bouaziz, Li, & Pauly, 2011). The technology currently requires a participant to "train" the software by scanning their face whilst holding various expressions before tracking, but this tool's ability to capture facial motion in real time without arduous application of facial markers is relatively new. Biocca states "Sensorimotor coordination is defined as the degree to which changes in body position correlate

immediately and naturally with appropriate changes in sensory feedback" and is one of the most important feedback channels responsible for presence as it parallels what we expect from our natural environment (Biocca, 1997, p. 10). Such real time autonomous facial motion capture should enable virtual world architects to better integrate users into the virtual environment.

Limitations of this Study

The technical challenges of creating realistic interactive works in 3D are many and varied. As this is an individual research project scope has been limited both technically and artistically. Immersive virtual worlds with Head Mount Display (HMD), head tracking of the Point of View (POV), and photorealistic skin textures were beyond the scope of this work. Future research would be well addressed by small teams with dedicated coders and 3D artists. Phenomenological and psychological explanations of feelings of embodiment in virtual worlds and uncanny aspects of CG characters have been addressed by other researches and will only be briefly commented on here (Biocca, 1997; Pollick, 2010; Steuer, 1992).

Methodology

This research utilises a practice-based research methodology to explore presence in facial motion capture driven avatars. The research is divided into two phases.

Phase One of this research experiments with a facially imitative avatar using "sketching" techniques and "rapid prototyping". It attempts to determine which key features of the avatars and their virtual environment are instrumental in eliciting different forms of presence for the avatar.

Phase Two of this study builds upon knowledge gained in Phase One and utilises the Iterative and Incremental Development (IID) method of software development (Larman & Basili, 2003). This is combined with a practice-based methodology to create an interactive digital work.

Seen holistically these two phases form a whole where inquiry and reflection feed

practice and creation of new knowledge. This is described by Schon as a process in which the

"… practitioner may surface and criticise his initial understanding of the phenomenon, construct

a new description of it, and test the new description by an on-the-spot experiment" (Schön, 1983,

p. 63).

Practice-based research acknowledges the importance of knowledge gained by the act of

creating an artefact, which in itself is seen as a form of knowledge. The knowledge and the

artefact are inseparable; one should directly reference the creative outcomes in order to

contextualise the knowledge (Candy, 2006). This research methodology blends well with the

"rapid prototyping" and IID methods of software development, using parallel processes of design

and development.

Phase One Process

The research was characterised by quick "sketches" and "rapid prototyping", producing a

series of prototypes that were "play-tested" within the extended Colab research community.

Faceshift facial motion capture software was used to capture participant expressions in order to

train software using the Kinect ("Kinect for Windows," n.d.) structured light scanner. After

"training" Faceshift was used in "tracking" mode, transmitting the participants Facial Animation

Parameter's (FAP's) in real time to Autodesk Maya using the Faceshift plugin for Maya and the

TCP/IP protocol. Maya and MEL scripts and MEL expressions were used to create, control and

render the virtual avatar, dynamic cloth and particles systems, and various props. Reflection on

the experience of both myself and my colleagues and analysis of the video footage captured gave

insight into which aspects of the simulation were contributing feelings of presence. Through a

series of iterations this process was used to develop the simulation by adapting algorithms to

affect results. A diet of readings of other work in the field informed design choices and suggested new directions.

Phase Two Process

Further to outcomes of the first series of prototypes in Phase One it was decided to create Face Echoes, a comment on selfies, using pre-scanned digital data and elements of AI to engage participants with the Duchenne smile; an invitation to interact with the avatar. Face Echoes uses scans derived from the training phase of a Faceshift motion capture session to build a database of expressions; three dimensional selfies with a smile, neutral, and a "not smile" expression requested of the participants. Toolkit for Sensing People in Spaces (TSPS) ("Toolkit for Sensing People in Spaces," n.d.) is then used in the performance to parse participants motions, a Java Processing application was written to receive the data via OSC, format in Maya's MEL commands, and send it on to Maya via Processing's net library. Maya then renders and plays back the selfies with selection criteria and transitions based on algorithms and fuzzy logic like function curves. The tuning of the work so as to engage passing viewers and the layering of complexity of responses to give the work a personality was achieved by an iterative process of adding functionality, testing on viewers in the Colab environment, further tuning, and then adding additional layers of response using the IID methodology of software development.

A large part of my ongoing practice consists of working in Maya to design systems that satisfy a visual end, in the case of a particle driven visual effect, or a user centric control system, in the case of a character rig. This is a creative technical endeavour, with experimentation of scripted code, system parameters, and overall system approach being applied. The advantage of a scripting language over a compiled language is the interactivity of the coding; in a compiled workflow the programmer codes, compiles, and tests the resulting simulation; whereas with

MEL scripting it is possible to change and adapt algorithms while observing the results in real time. A scripting workflow enables playful programming; a key component of my practice allowing a iterative development process that incorporates reflection of results observed into the development cycle.

Literature Review

Structure of Review

This review is organised around the following topics central to our perception of CG avatars. These will be briefly presented and then discussed in the body of the review.

CG Avatars - Avatars have deep roots in our society and mythos from the story of Narcissus in which the young hunter becomes entranced with his own image in a pool of water, to the popularity of the selfie stick and selfie drones the fascination with our own face permeates the history of cultures.

The Uncanny - The uncanny is the achilles heel of CG characters and has been the subject of research recently; technology and expectations of photorealism are driving CG imagery forwards but the specific aspects that create uncanny responses are still unclear.

Selfies - Many commentators suggest that we are living in an era of unparalleled narcissism with the rise of the selfie and others say the selfie is a continuation of self portraiture even being on the brink of a new genre of the art form; sefies and avatars are seen by this research as being driven by the same motivators.

Presence in Virtual Environments - Presence; the sense of "being" in a virtual or mediated environment, is a critical design component of virtual reality simulations and the existing literature on types and causes of presence is key to development of CG avatars.

Facial Action Coding System - Ekman's work on facial expression deconstruction and on discerning basic emotions from these expressions is a key element of this research and is utilised in CG character animation tools and motion capture softwares.

Facial Feedback Hypotheses - The Facial Feedback Hypotheses (FFH) is the motivation for the creation of the simulation "Flowers" detailed in this paper, although this study does not set out to provide a proof for the FFH.

Human Computer Interaction - Human Computer Interaction studies are recently using our facial expressions as one means to bridge the gap between our emotional state and the manner in which software responds to us and recent work in this field is important to position this research.

Discussion

Avatars and Narcissism

Cleland discusses CG avatars in context of our long standing fascination with our mediated image in "Image Avatars: Self-other Encounters in a Mediated World" (Cleland, 2008). She describes the history of our fascination for the mirror; from the myth of narcissus, through painting, photography and the moving image.

The doubling of observer into "self and other" that results from contemplation of our image is at the heart of telepresence (Cleland, 2008, p. 47). Cleland identifies the face as the most compelling body part contemplated in the mirror image as it is not seen directly, only in a mediated form. She discusses how in the mediated mirrors of portraiture, photography and film there is a temporal displacement of the image; only live video installations and mirrors provide a real-time experience.

Cleland also speaks to the "Uncanny Valley" when she discusses how the mediated image of one's self has always been uncanny. Photographs cheat death and embalm the memory (Barthes as cited in Cleland, 2008) and film scholars have referred to moving images as phantoms in a new kind of mirror (Metz as cited in Cleland, 2008).

Cleland argues that image avatars have evolved over the past one hundred and fifty years and contemporary technologies have brought us novel mediated experiences of ourselves (Cleland, 2006). The mirror, the video screen, and the motion driven avatar all offer us new views of our body that promote the oscillation of our viewpoint between "here", in our body and "there", where our body seems to be (Biocca, 1997).

Real time avatars have to date been constrained by polygon budgets and have suffered from a small selection of pre-modelled facials shapes to represent emotional states. Biocca argues that a higher resolution of body sensors increases sensory engagement and is responsible for progressive embodiment. He defines progressive embodiment as "the steadily advancing immersion of sensorimotor channels to computer interfaces through a tighter and more pervasive coupling of the body to interface sensors and displays" (Biocca, 1997, p. 5). As the face is our most expressive communication organ it stands to reason that developments in motion captured avatar facial fidelity will increase the communication efficacy of our avatars.

The Uncanny

Perhaps the first mention of the uncanny effect of artefacts appearing near human was by the german psychiatrist E. Jentsch. In his essay "On the psychology of the uncanny", written in 1906, he reasons that the most consistent method of causing an uncanny feeling are situations where "doubts whether an apparently animate being is really alive; or conversely, whether a lifeless object might not be in fact animate" (Jentsch, 1997, p. 6).

Thirteen years later S. Freud wrote "The Uncanny" defining his subject as: "The uncanny is that class of the frightening which leads back to what is known of old and long familiar" (Freud, Strachey, Cixous, & Dennomé, 1976, p. 27). In "The Uncanny" Freud also alludes to the fascination with our own image. He discusses narcissism and mirrors, arguing our first double or avatar was our immortal self; an escape from the inevitability of death. But when our self image becomes real, and animate: "The double has become a thing of terror, just as, after the collapse of their religion, the gods turned into demons" (Freud et al., 1976, p. 38).

"The Uncanny Valley" by Masahiro Mori, robotics professor at the Tokyo Institute of Technology, was published over 40 years ago (Mori, MacDorman, & Kageki, 2012). This journal article was a thought experiment that hypothesised the negative and positive affinities that we would feel towards artefacts or representations of humans as they progressively become more humanlike and Mori describes the point at which artefacts become terrifying as Freud had described.

Mori presents this information in the form of a graph in which human likeness is mapped against natural liking for an artefact or representation.
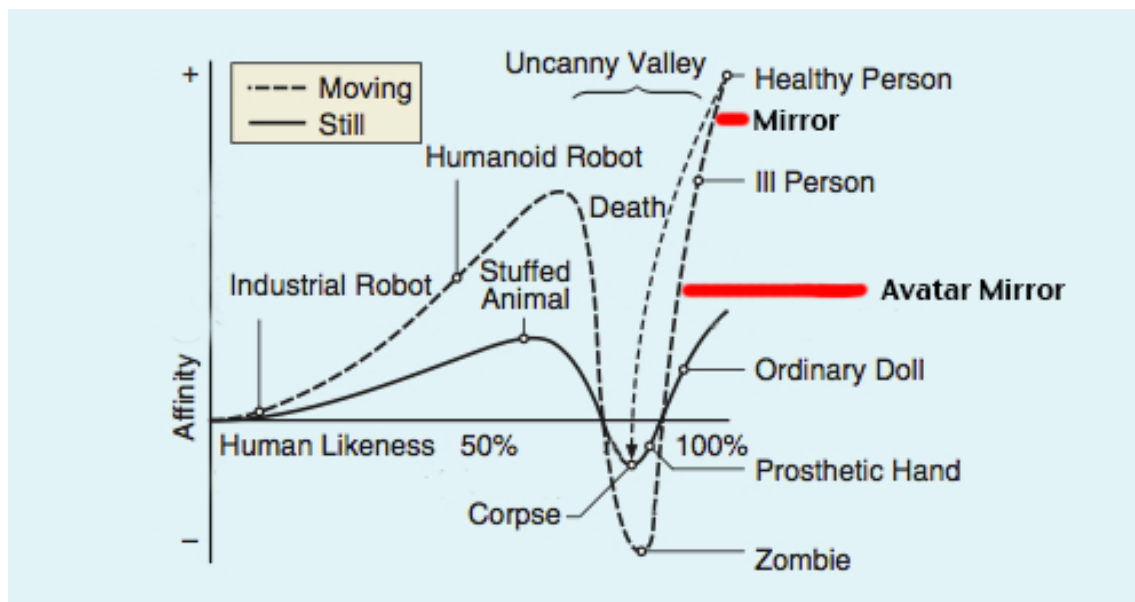
Figure 1. The Uncanny Valley. Based on image by Masahiro Mori and Karl MacDorman, ©
2012 IEEE. Reprinted with permission.

Figure 1 is a diagrammatic representation of his hypothesis, with values of affinity
plotted against closeness of representation of human likeness. There are two plots, one for a
moving representation and one for a still representation. According to Mori "creating an artificial
human is itself one of the objectives of robotics" (Mori et al., 2012, p. 1), but the paper was not
discussed in academic circles until CG animated characters began to exhibit uncanny
characteristics. The positions of mirrors and avatar mirrors have been added to the "moving"
curve to illustrate two postitions this research would take.

Maddock et al have presented a system to animate a face using motion capture data as
well as extrapolating facial wrinkles to enhance realism (Maddock, Steve, Edge, James, &
Sanchez, Manuel, 2000). This system uses a weighted pose based expression set to animate the
face from the motion capture data while centring the shape on the dominate viseme to more
closely replicate the dominance of shapes that occur during rapid speech where visemes blend
together. Their conclusions state that in order to avoid uncanniness perceived in "embodied
conversational agents in interfaces, and for agents or avatars in virtual environments, movement
and behaviour are more important than photorealism" (Maddock, Steve et al., 2000). The study
was done using highly detailed characters but summarises "there seems to be no evidence as to
how realistic facial movement such as correct mouth shape and wrinkles has to be" (Maddock,
Steve et al., 2000, p. 3). Maddock argues that the overall character behaviour is more important
than facial wrinkle details. "The character must be seen to react in an intelligent and human-like
manner" (Maddock, Steve et al., 2000, p. 3).

The Uncanny Valley hypothesis was recently tested against hand animated CG character pre-rendered facial expressions (Tinwell et al., 2011) . Lack of upper facial cues were found to increase uncanniness for all of the six basic emotions; anger, disgust, fear, happiness, sadness, and surprise (Paul Ekman, 1992a), but the authors found that upper facial cues were not contributing to uncanniness in conveying the emotions anger and happiness (Tinwell et al., 2011). Happiness was still rated lowest in human-likeness of all the six emotions tested, attributed by the authors as possibly because the lack of spontaneity of the modelled actor's smile and the smile being perceived as false (Tinwell et al., 2011). The study did not specifically use the Duchenne smile (Soussignan, 2002) as did the "Flowers" simulation detailed in the "Practice" section of this paper. The results of this study have been valuable in the design of the current research although the data does not include use of motion captured CG characters.

Recent advances in cinematic technology have enabled film makers to achieve CG characters that bridge the uncanny valley. Notably uncanny human-like characters in "Final Fantasy" 2001 and "Polar Express" 2004 served as proof that the uncanny valley did indeed exist. "The Curious Case of Benjamin Button" 2008 was the first CG double where the animated character was indistinguishable from the real actor; Brad Pitt. "I'd swear it was Brad Pitt right there on the screen—and I'm an evil-eyed, annoying computer-graphics expert" (Alvy Ray Smith as quoted by Perry, 2014, p. 2). The caveat to this achievement being images that make up one twenty-fourth of a second of movie can only be produced by multi-core racked processors working in concert for days, and many man hours of post-processing by graphic artists; real time performance at this level is not yet achievable.

Current works in cinematic films have made advances in crossing the uncanny valley, and real time game developers are not far behind. Real time avatars can have different feature

priorities and use design choices driven by the goal of communication of emotional states. As

Tinwell et al. conclude, "designers wishing to successfully depict anger or happiness in their

character need not invest substantial amounts of time in embedding upper facial expressivity,

especially so if lower facial features and prosody support recognition" (Tinwell et al., 2011, p.

8).

Selfies

Selfies are a contemporary manifestation of our fascination with the image of our face,

the most expressive part of our body but one of the few we cannot see without technological aid.

The first selfies could be said to be early self portraitures painted from mirrors. Saltz describes as

a "proto-selfie" Parmigianino's 1523–24 Self-Portrait in a Convex Mirror (Saltz, 2014). The art

of self portraiture was enabled by two technologies; inexpensively mass produced mirrors and

the advent of photography.

"We are hard-wired to respond to faces, it's unconscious. Our brains process visuals

faster, and we are more engaged when we see faces. If you're looking at a whole page of photos,

the ones you will notice are the close-ups and selfies" (P. Rutledge as quoted by Wortham,

2013).

The word "selfie" was declared word of the year by the Oxford English Dictionary and

included in the online version in 2013. Although the form has generated huge interest online;

research and analysis of selfies in the context of the media that they are published in are

relatively sparse.

Selfies have started to evolve into new sub-genres. Selfies taken in mirrors instead of at

arms length can be said to be a special subset of the genre (Saltz, 2014). This is a telling

evolution of the form, where technology uses its predecessors to reinvent itself. Selfies taken of

ones avatar are again a remediation of the form, with there being an entire website devoted to the video game "Grand Theft Auto" selfies where game users take pictures of their avatars/characters in game (GTA 5 Selfies, n.d.).

Time lapse selfies are the only animated mutation of the form. In September 2014 Instagram added time-lapse photography to their application to compete with Apples IOS8 features recently released and called it "selfielapse". Time lapse is a specialised type of temporal image change, using a moment from each day to create movement with the interest to the viewer being possibly the ageing of the subject, an accelerated sense of time.

The selfie can be shown to be an amalgamation of our fascination with our face, the emotional content of our expression, and new mobile technology. It shows the extent to which we embrace new technology to better our view of our faces and communicate our feelings; much the same way as in which mass production of mirrors enabled self-portraiture. The facially imitative avatar stands at the cusp of being reinvented as another technologically enabled mutation of the self portrait.

"Marina Galperina, who with fellow curator Kyle Chayka presented the National #Selfie Portrait Gallery, says, "It's less about narcissism—narcissism is so lonely!—and it's more about being your own digital avatar" (Galperina as quoted by Saltz, 2014, p. 3).

Embodiment in Virtual Environments

In "The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments" Biocca discusses designer challenges and choices enabling embodiment in immersive virtual worlds (Biocca, 1997). Biocca defines the goals of progressive embodiment as continuously increased presence in virtual worlds. He writes that "Users experiencing presence report having a compelling sense of being in a mediated space other than where their physical body is located"

(Biocca, 1997, p. 12). He reinforces Freud (Freud et al., 1976), arguing that our ancient desire to experience transcendence and the "essential copy" of ourselves are driving the pursuit of mediated presence (Biocca, 1997, p. 13).

Biocca argues the cyborg's dilemma is; "the more natural the interface, the more we become 'unnatural', the more we become cyborgs" (Biocca, 1997, p. 23). Cyborgs exist in the depth of Mori's uncanny valley, at once familiar yet belonging to the strange and unknown. Biocca goes on to ask "what is 'natural' about our relationship to our technology" (Biocca, 1997, p. 23). He says "We tend to think of technology as something alien, not a reflection of ourselves" (Biocca, 1997, p. 23). The marriage of technology and ourselves in our mirrored avatar is an impediment to us accepting our embodiment in the avatar; our reluctance to become cyborg.

Research in the field finds that presence in virtual enviroments is a multi-faceted phenomenon (Kim & Biocca, 1997), and various forms of presence have been postulated. Presence and embodiment was recently reviewed by Schultze (2010) where he identifies six different forms of presence in virtual reality identified in previous research with the one that best characterises embodiment in a facially imitative avatar being "self-presence", defined as "the users' mental model of themselves inside the virtual world" (Biocca as cited by Schultze, 2010, p. 6).

Biocca also speaks to the idea of social presence in virtual enviroments; where communication between two intelligent beings is the key element. "The ideal of the face-to-face interaction is the background against which comparisons are made for all technologies that link two humans together" (Biocca, 1997, p. 18). He also argues that artificial intelligence can take the place of of a human agent and that "Users may be fooled by convincing morphology and believe an artificially intelligent agent is really a humanly directed avatar" (Biocca, 1997, p. 19).

These are key concepts in the design of "Face Echoes", where engagement with a participant using a realistic data scanned smile helps to establish authenticity.

Biocca's arguments indicate real time facial motion capture should result in enhanced feelings of embodiment in virtual environments that are responsive to this data because of increased fidelity of the motor channels captured. This appears to have been borne out by the results of this study.

Facial Action Coding System

Ekman's long research into facial expressions and emotional states has proven how closely the two are linked (Paul Ekman, 1992b; Paul Ekman & Friesen, 2003; Paul Ekman, 1992a; Paul Ekman, Friesen, & Levenson, 1990). His research and publications build upon work dating as far back as 1872 citing Darwin's "The Expression of the Emotions in Man and Animals" as the basis for his current theories on emotional families and the evolutionary survival origins of same (Darwin, 1998).

Ekman and Friesen developed the Facial Action Coding System (FACS), (see Paul Ekman & Friesen, 2003), which is a system that uses Action Units (AU) to describe anatomical muscle changes in facial expressions, giving trained observers insights into emotional states based on expressions and micro-expressions. FACS has been adapted to the Facial Animation Parameter description language (FAP), for use in computer animation and FAP was included in the MPEG-4 media bundle in 1998 (Pandzic & Forchheimer, 2003). FAP's can be used to control a CG character's facial expressions to create the illusion of speech and emotion.

Faceshift ("Faceshift," n.d.) is a facial motion capture tool that outputs FAP's to describe facial expressions for CG animation purposes. One of the advantages of this mapping is that it is easily repurposed across different facial topologies, or even species (see Sagar, 2006). Other

recent implementations of FAP's as a description for facial motion data are by Affdex ("Facial

Expression Analysis Recognition-Affectiva," n.d.) and Emotient, ("Emotient - Facial Expression

Recognition Software," n.d.), both marketing tools that "read[s] facial expressions to measure the

emotional connection people have with advertising, brands and media" ("Facial Expression

Analysis Recognition-Affectiva," n.d.).

The relationship between Ekman's work on the emotional context of various

combinations of FAC's and the expressions captured by Faceshift expressed in FAP's serves as a

useful mapping for doing research in this area. The workflow in this research was able to utilize

this mapping without the overhead of programming ground up detection algorithms.

Facial Feedback Hypothesis

Ekman's research on the Facial Feedback Hypothesis (FFH) uses FACS for a standard to

manipulate and evaluate subject's facial expressions during experiments. This research uses

FAP's, a derivative of FACS, to drive the CG avatar with motion capture data.

One of the models for facial feedback is the sufficiency hypothesis which holds that

facial expression on its own can generate an emotional response. This has been borne out by

research (P. Ekman, Levenson, & Friesen, 1983).

Ekman (P. Ekman et al., 1983), Strack (Strack, Martin, & Stepper, 1988), and other

researches (Soussignan, 2002), use pencils held in subject's mouths to achieve a Duchenne smile

(Soussignan, 2002) or a lips together sulky condition in order to establish a proof for the

sufficiency hypothesis. See Figure 2 showing "sulky" condition and Duchenne smile achieved by

asking participants to touch or not touch the pencil with their lips. Recent research in this area

has used electromyography (EMG) measurements to induce subjects to hold a facial expression

whilst their responses to positive and negative images were recorded in a questionnaire (Dimberg

& Söderkvist, 2011). The Duchenne smile is a critical human expression that crosses cultural

boundaries and may enable avatar designers to bridge the uncanny valley.

Figure 2. Showing sulky condition "b" and lips corner pulling and cheeks raising (Duchenne

smile) "d". Based on image retrieved from

http://www.researchgate.net.ezproxy.aut.ac.nz/profile/Robert_Soussignan/publication/10630044

_Duchenne_smile_emotional_experience_and_autonomic_reactivity_a_test_of_the_facial_feedb

ack_hypothesis/links/0912f50f9501286a79000000.pdf © 2002 IEEE. Reprinted with permission.

This research suggests possibilities for using facial motion capture as a measuring tool or

in a feedback mechanism to induce specific expressions. This would be a useful methodology to

test the FFH and to regulate emotion.

Human Computer Interaction (HCI)

HCI research is attempting to bridge the gap between computer and user; creating new

interfaces to better our experience. In order for us to better use our technology, or become more

cyborg (Biocca, 1997), one interface that will improve our experience is to have our machines

better recognise our emotional states.

Ekman's work provides developers with the mappings to bridge from emotional family (Paul Ekman, 1992) to FAC, and then to FAP using data input from motion capture and animation output to CG avatars. A contemporary example of on-going research in this area is MIT labs facial expression recognition systems being designed to bring awareness to vehicle drivers of their emotional state in an effort to promote a safer driving experience ("AutoEmotive," n.d.).

The depth sensors cameras in the Microsoft Kinect were marketed aggressively by Microsoft in their recent Xbox one as a revolutionary way in which we could interact with our computer based on skeletal and gestural controls. They were to be bundled with every console sold, but they failed to gain commercial traction and did not deliver any compelling content so Microsoft "un-bundled" the depth camera. Intel "RealSense" technology is recently being promoted by Intel and has the promise of delivering a depth enabled facial and hand gesture motion capture solution to every laptop device ("Intel® RealSenseTM Technology," n.d.).

The general dearth of compelling content using this technology even though the "big names" are still commited to developing HCI interfaces using facial expressions and hand gestures positions research in this field as essential and timely.

Practice

Phase One: "Sketching"

In order to investigate the research question posed in this study IID and rapid prototyping were used in Phase One to develop a series of sketches or prototypes alongside a literary investigation of the field to determine what factors enhanced embodiment in a facially imitative

avatar. Uncanniness was identified as a negative outcome and prototypes that exhibited this response were not developed further.

The mirrored countenance of the computer monitor as metaphor for a mirror for our face is a comfortable medium, one that we have been used to since we were infants. This is a good starting point for experiments with the avatar, avoiding the uncanniness that more immersive mediations such as HMD have induced as a distraction to presence. Computer monitors were used to display the participants video captured face, the resulting Faceshift avatar full head representation, and the Maya simulation.

The technical setup for Phase One used a Kinect scanner for structured light depth scanning and video capture, Faceshift facial motion capture software for interpreting facial expressions and their resulting FAP's, the Faceshift for Maya plugin delivering a real time stream of FAP's data into Maya, and Maya for virtual world behaviours and 3D rendering. Apple Quicktime Player was used to screen capture the performance and save a .mov file for documentation.

Video documentation of all simulations created in Phase One is viewable at https://youtu.be/xYZFPPR6pYM.

"Weather Face"

Facial Expressions controlling a Virtual Environment using FAP's associated with Happy, Sad, and a "blowing" expression.

"Weather Face" is a simulation driven by FAP's associated with the emotional families Happy, and Sad (Paul Ekman & Friesen, 1977). Happy makes the "sun" get larger, and begins to melt any "snow" on the ground. "Sad" grows the "cloud" shape and puts it in front of the "sun".

A mouth funnel shape emits Maya particles into the world and the particles fall as "snow",

lingering on the ground until the "sun" melts them. See Figures 3 and 4.
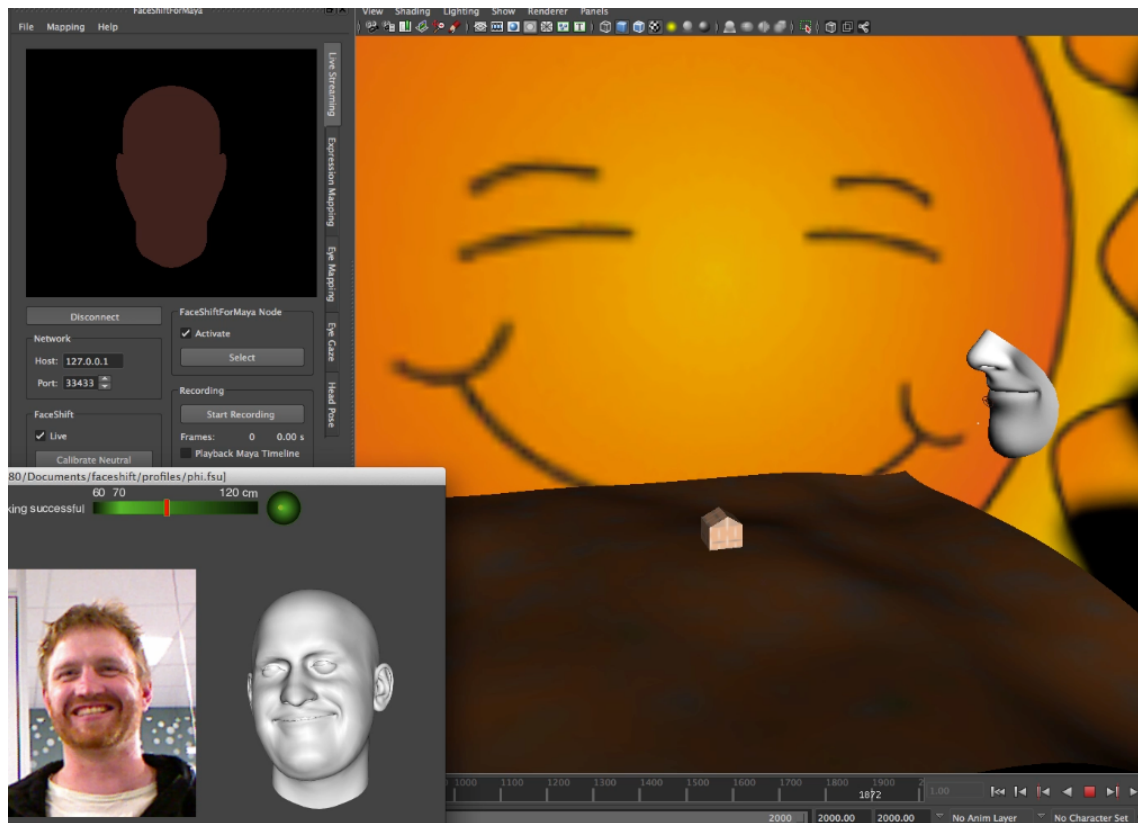
Figure 3. "Weather Face" showing smile with enlarged "sun".
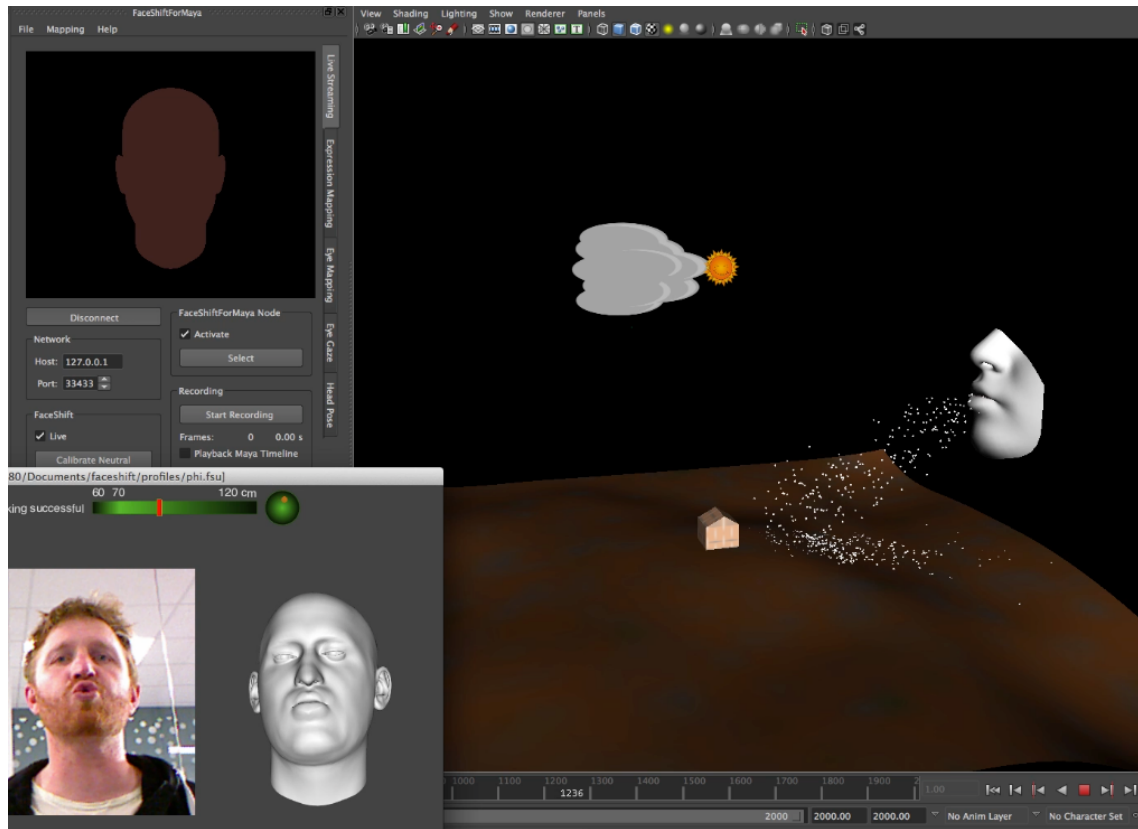
Figure 4. "Weather Face" showing blowing "snow", "cloud", "sun".

The simulation was attuned to the various expressions by capturing myself making the

relevant expression in Faceshift tracking mode, recording the levels of the FAP's, and then

writing an algorithm to effect the environment parameters based on the levels of FAP's.

Participants described the experience as "busy", but the work showed that using facial

expressions to expel virtual objects had a high level of embodiment, with participants regularly

blowing air out of their real mouths in an attempt to propel virtual objects. The other

manipulations of the "sun" and "rain" were interesting and engaging to participants, but users

found them difficult to control.

Early in the design of "Weather Face" I had user feedback that the eyes in the model

were "creepy" and "distracting". The eyes in the model were simple from a cinematic and

anatomical point of view, only having the ability to track the participants rotation and lacking

basic anatomy such as lacrimal caruncle and deforming eyelids, both crucial to a convincing morphology.

Based on this feedback the model was simplified to only concentrate on lower face expressions, reducing the number of FAP's in use from 68 to 23. The model was textured in porcelain white, reducing detail that could detract as the user evaluates the "photorealism" of the skin texture. These changes forced the viewer to see the facial expressions without bias based on real time render compromises. This choice was based on results from Tinwell et al. where that study found that "A lack of upper facial cues during speech did not influence perceived uncanniness for the emotion happiness" (Tinwell et al., 2011, p. 7).

"WeatherFace" is the only work that is located in a virtual environment with props, ground-plane, sense of scale. Participants did not report a strong sense of presence in "WeatherFace". The act of creating virtual material with the mouth in the CG environment was shown to be more compelling than seeing an imitative avatar in an environment.

"Blow"

Breath modality and effect on virtual objects.

"Blow" was created to focus on just the act of expelling virtual particulate matter using Maya particles. See figure 5.
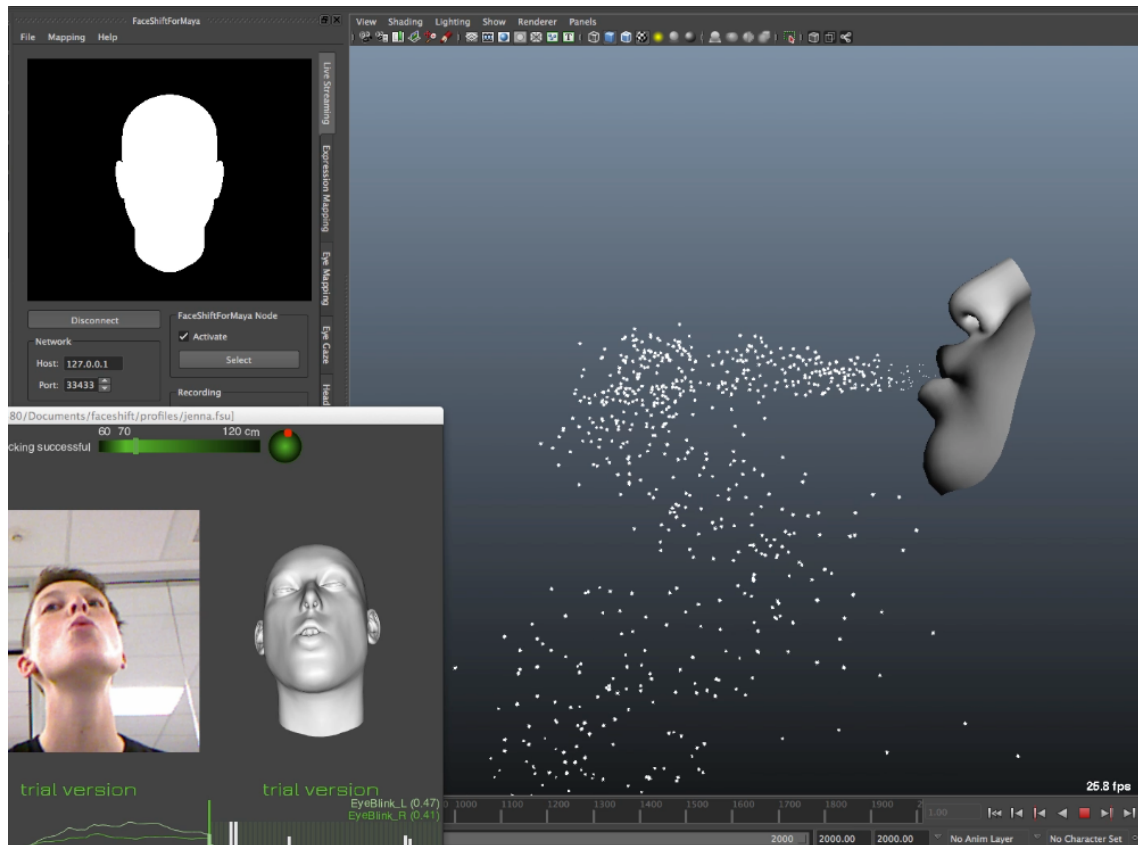
Figure 5. "Blow" showing funnel shape triggering particle emission rate and speed.

There are five mouth shapes that control the simulation in concert. The shapes are not indicative of an emotional state but required the user to assume a functional state; how we blow out a candle, or exhale in a forceful or directed manner.

To ascertain which shapes comprised an exhalation the facial shape appropriate to blow out a candle at a distance was held, complete with a forceful exhalation, and the resultant shapes captured with the facial motion capture software. The resulting FAP's transmitted from Faceshift were studied to develop the algorithms driving the simulation. The avatar's head rotates with the participants motions enabling directional control of the stream of particles. See figure 5.

Users consistently blew air out of their mouths in an attempt to produce more virtual material in the simulation. This is indicative of strong feelings of embodiment, a connection of the virtual world to the user's body schema. Replacing the participants breath with virtual snow

bridged the two modes; facial expression and the highly reflexive act of breathing but did not result in the typical oscillation of viewpoint that Biocca speaks of in virtual worlds when the user in embodied in the world, defined as self-presence (Biocca, 1997).

"Bubble" and "Chew"

Embodiment of a virtual object by manipulation through mouth shapes.

Manipulating virtual objects with lower face shapes promoted feelings of embodiment in the previous sketches so I produced a series that enabled the user to manipulate a virtual object in their mouths; "Bubble" and "Chew". "Bubble" had the object quickly expelled from the mouth, being able to be shaped with the lips. See figure 6. "Chew" promoted "playing" with the virtual blob with the lips and offered the opportunity to continue shaping the virtual matter. See figure 7. "Bubble" and "Chew" used Maya's dynamic cloth solver to simulate the collisions, deformation physics, gravity, and directional forces that expelled the cloth object. The simulation was restarted whenever the virtual blob exceeded a predetermined threshold from the avatar.

Figure 6. "Bubble" showing transient manipulation and shaping of virtual object using mouth shapes.
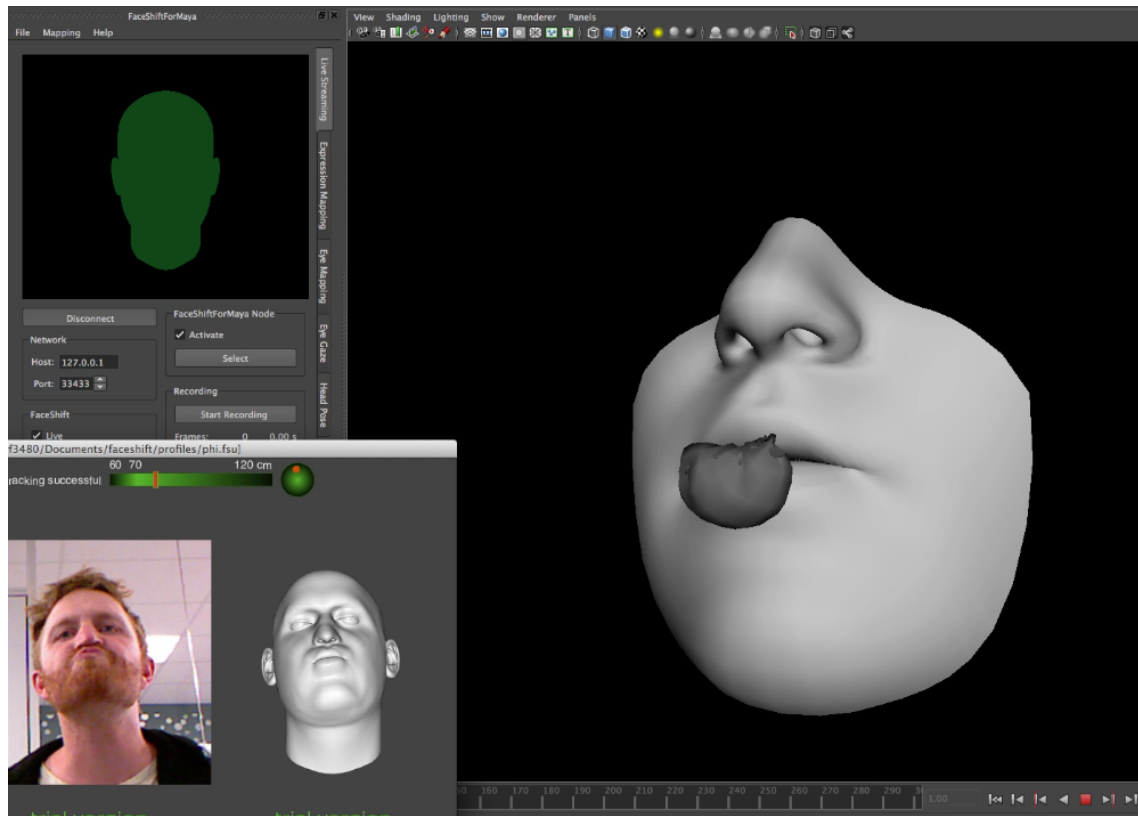
Figure 7. "Chew" showing persistant manipulation and shaping of virtual object using mouth shapes.

Users found these works engaging, trying to keep the virtual object from leaving their mouths, bouncing the object with their lips, and mouthing the object into new shapes. "Chew" was created on reflection of these comments, allowing the virtual object to be more easily contained. Comments from participants were that the virtual gum or ball was felt embodied inside the mouth in some cases. This would be consistant with descriptions of self-presence.

"Flowers"

Facial Feedback Hypothesis experiment with the goal of affecting a virtual object

outside the body.

Flowers is a work designed to leverage the facial feedback hypothesis to generate

positive feedback in users. See figure 8 showing Duchenne smile triggering progressive growth

of "roses" with feedback loop.

I created a variation of the classic study by Strack (Strack et al., 1988), whereby the

subject's emotional state is tested using cartoons and pictures whilst holding facial muscles

rigidly. Later studies by Soussignan specifically activated a Duchenne smile (Paul Ekman,

1992b) and found that participants reported "more positive experience when pleasant scenes and

cartoons were viewed" (Soussignan, 2002, p. 1).
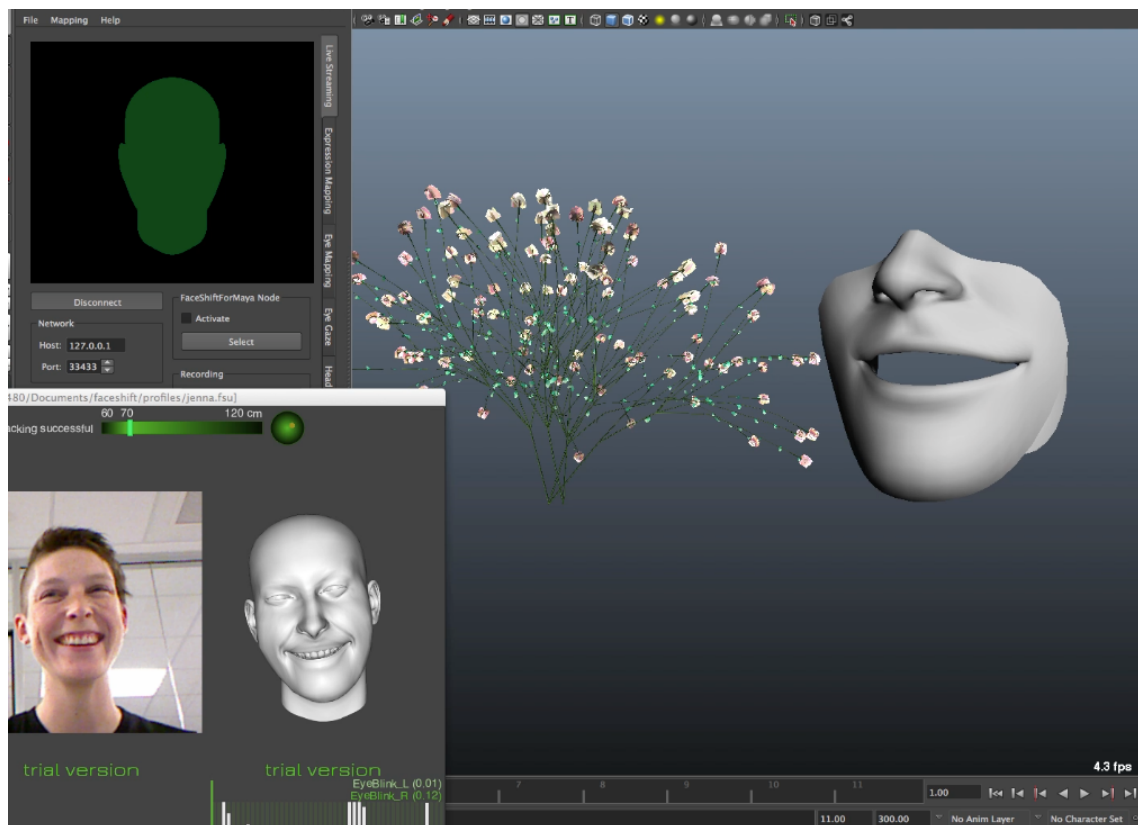
Figure 8. "Flowers" showing Duchenne smile triggering progressive growth of "roses" with feedback loop.

The Flowers simulation uses the levels of Action Unit (AU) 6 (orbicularis oculi cheek raising), and AU12 (zygomaticus major lips corner pulling), that identify the Duchenne smile, to cause a Maya Paint Effects rose to grow. Increased levels of these AU's cause more branches and rose flowers to appear depending on algorythms MEL scripted in the Maya scene. Participants were not given instructions or guidance but encouraged to approach the simulation playfully. After some moments users found the growth of the flowers was contingent on them smiling and began to experiment with holding a smile to make the flowers grow larger. This was successful in causing participants to keep smiling, and sometimes begin to laugh; perhaps creating a cascading feedback loop. Users also reported very low latency and high levels of responsiveness to their emotional state in the simulation. There was no presence reported by participants engaging in this simulation.

Phase Two: "Face Echoes"

"Face Echoes" is an interactive work that provides a mirror with a different perspective, mirroring not the participant but echoes from temporal and emotional spaces beyond the installation. It was exhibited as part of Digital Art Live's "Emerging Pixels #4" at Aotea Centre, Auckland, in December of 2014. See figures 9 and 10.

Figure 9. "Face Echoes" showing intial smile behaviour engaging participants. Retreived

from https://vimeo.com/119073497. Reprinted with permission.
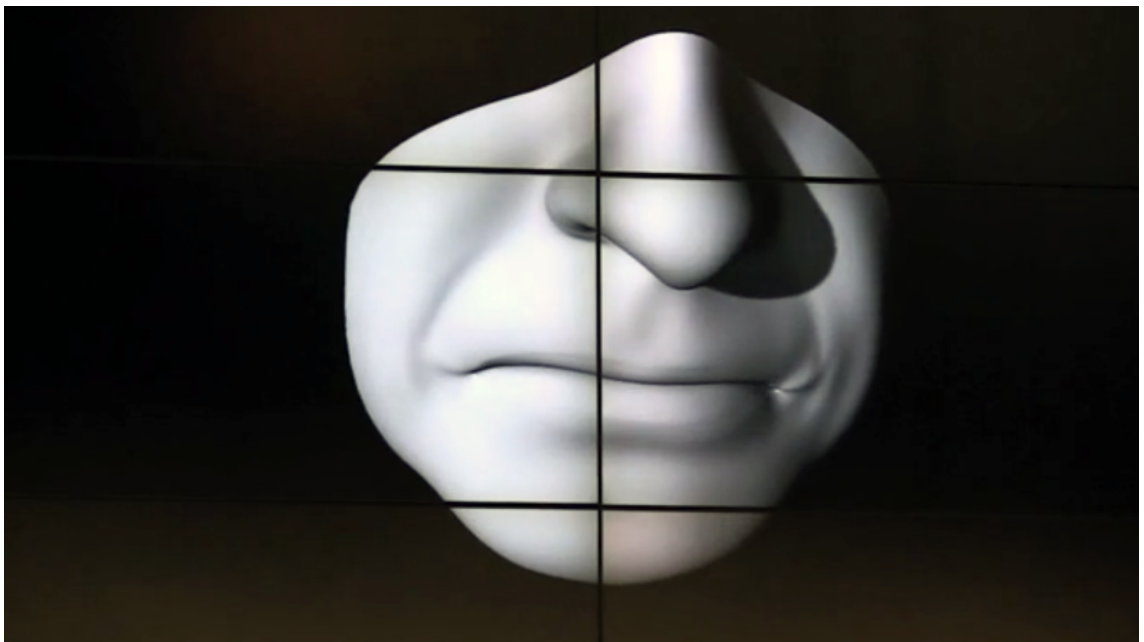
Figure 10. "Face Echoes" showing "not smile" due to participant's velocity and time

spent in capture space. Retreived from https://vimeo.com/119073497. Reprinted with

permission.

"Face Echoes" is imagined as a collection of selfies that peer out from the virtual mirror,
attempting to engage passing strangers. The selfies are acquired by digitally capturing fellow
researcher's faces at Colab AUT using a structured light scanner while they give to the work
expressions of joy, un-happiness, and neutrality. Based on the participants movements, time
spent in the detection space, and the number of participants in the space, algorithms decide
which type of expression to display. The facial expressions morph from one face to another
creating a new, virtual personage, transient in nature. This parallels the transient nature of facial
expressions and our emotional states and comments on the static fixing of our face by the
medium of the selfie. "Face Echoes" speaks to multiple themes identified in Phase One of this
study and in the review of literature in the field. These will be addressed in the following
sections.

Uncanniness

"Face Echoes" created uncanny responses reflected in comments that it was creepy at
times by some viewers. The Duchenne smiles that "Face Echoes" presents upon recognising a
participant in the space were well received, but as the user spent more time in the capture space
and began to display neutral and then the "not-smile" expression users commented that they were
disturbed by the simulation. This confirms results of the study by Tinwell et al. where it was
found that "A lack of upper facial cues during speech did not influence perceived uncanniness
for the emotion happiness. Indeed, the virtual character was regarded as slightly more human-
like in the lack condition" (see Tinwell et al., 2011, p. 7).

The Mediated Face

Media enabled by emergent technology is now extending our most expressive organ, broadcasting our captured state and repeating our emotive state again and again as echoes. Our face can express emotions to others at a great distance. We are very adept at interpreting these signs of intent. The selfie, a contemporary mirror, reflects our mediated expressions and the emotions that they convey.

Beginning with the telephone, we have become normalised toward projecting a mediated version of ourselves, a transported mirror. Now when we gaze intently at our "personal device" on arm outstretched, we have produced our avatar; the face and emotive state we display for the camera are projected on a web of light around the world; we stand behind the avatar. It comes as no surprise Facebook is named according to this compulsion; it lets us gaze upon our own visage as we browse the faces of others from behind the mirror of the computer display.

The Animated Morph

The morph itself is an intermingling of facial intent, creating for a moment in time a new visage, a "gutter space" of the bits between the selfies. Animated selfies have yet to be "discovered" by the masses; "Face Echoes" explores this space with a selfie that engages with the viewer with people tracking, watching participants as they watch the avatar change its shape and orientation. The avatar is for a moment not a representation of one person, but a person between, a new being only present in cyberspace. This collation of a smile and a frown to create a new, transient being, only to shift again to the permanent record of a "real person" is unsettling to the viewer, not knowing how to react to an illusionary expression.

Focal length and Low Key Lighting

"Face Echoes" employs low key lighting to juxtapose the image against the selfie's natural environment. Selfies are normally taken in high key natural lighting due to limitations of the technology. Camera shots that include flash fill or lighting from the camera position are rarely flattering and not often seen in selfies. One of the common artistic devices that characterise a selfie is high key natural lighting with natural fill. "Face Echoes" juxtaposes these structures with dramatic low key lighting utilising high subject key to fill rations and high subject to background ratios. The face is alone in a dark landscape, bringing the loneliness of narcissism to bear on the viewer. The lighting accentuates facial features, the negative space grows and denies a frame of reference for the mirror of the screen.

Selfies are usually shot with a large focal length the I-Phone 5 rear camera has a focal length of 31mm, whereas "Face Echoes" uses a focal length of 22mm to capture the image in the work. This short focal length is used in the virtual camera to emphasis the foreground element, distorting the facial expressions in an unusual way.

Interactivity and Face Echoes

Participants in "Face Echoes" try to engage the work by making "faces", posturing, wondering if the face is actually responding to their facial expressions. In human to human interaction imitative expressions are normal; we smile, we expect a smile back. The lack of intelligent human response is not immediately evident in "Face Echoes", forcing participants to try harder to communicate. The urge to communicate is strong and the offer of friendship that a smile promises does not go un-noticed by the viewer. This response is indicative of social presence, communication between one intelligence and another, even if one party is Artificial Intelligence driven (Biocca, 1997, p. 18).

Surveillance

Surveillance is alluded to in "Face Echoes"; even without eyes the face rotating to follow the viewer and reacting quickly to new viewers makes us wonder at the intent of the intelligence behind the face in the work. With the current archiving of meta data and facial recognition software cataloguing huge databases of faces from public surveillance cameras viewers of "Face Echoes" contemplate whether their faces have indeed been "captured" by the work for some nefarious intent. Having no eye gaze to give a direction to the avatar's line of sight does not seem to impair the feeling of being "watched", perhaps enhancing the feeling of surveillance. The face that follows ones movements is evidence enough of intent, even though the intent is not immediately obvious.

Mirrors

The screen and the mirror have much in common; an immediacy, a window into another world. "Face Echoes" mimics some of a mirror's characteristics; our image turning as we move to face us always, no temporal displacement, our image receding as we move away. The void that the avatar inhabits is not mirror like. The realisation that the avatar is not mimicking the facial intent of the viewer, the reversed behaviour unlike a mirror that has the face observing the viewer from oblique angles juxtapose a mirror's behaviour against the avatar's point of view.

Technical Specifications

"Face Echoes" was built using Toolkit for Sensing People in Spaces (TSPS), the Processing Java development environment ("Processing.org," n.d.), and Autodesk Maya. Applescript was used extensively in automated start and shutdown procedures to enable Maya, a

timeline based animation software, to play over four million frames a day, every day for a

month.

"Face Echoes" inputs depth image data captured live from the Kinect structured light

scanner. This is then parsed in TSPS and sent to a Java application using the TSPS Processing

library.

Java code was developed in Processing using example code from emil@polyzaar.com

and TSPS forums as a starting point. This code parses the incoming TSPS data and tracks only

the person that was in the scene longest. It then sends the positional, velocity, and depth data

from the sensor to Maya, as well as a current person id that changes when the oldest, (longest

time in sensed space), person in the scene is changed. The number of people in the scene as well

as the distance from the Kinect camera that the subject is is also sent to Maya using Processing's

.net library.

In Maya there are five expressions on locators that break a "rule" of expressions by

directly setting and attributes, usually not done as it "breaks" the scene graph refresh. Many

attributes including the model refresh in viewport without the timeline playing. This is a very

interactive form of programming as some expressions will not parse if the timeline is not on

frame one where global variables are declared in some scripts. Maya's scene graphs do not work

well with multiple simultaneous events arriving at the same node, this leads to indeterminate

results. The likelihood of simultaneous multiple events was minimised by using a "latch" system

inspired by Massive software's motion library latches ("Massive Software – Simulating Life,"

n.d.). When a latch goes "high", or to the value of 1 on a certain part of the system, no incoming

events are accepted until the the system finishes the requested procedure and latch is once again

set "low", or 0.

Although Maya is not generally used as an interactive engine, it will accept eight million frames on the timeline. A script node in the scene starts the simulation playing in "Interactive Mode", a type of playback designed for working with Maya's several dynamic engines.

The following Maya scripts are attached to various scene objects and form the bulk of the logic system.

Info Exposes position, velocity and personid data, swaps positional data between the current moving target and next target. Sets flag of which target is being animated.

Constraint Does an aim constraint swap between the last active target and the next target. This is animated and the duration is exposed for further development.

Blend Parses the three blendshape families and the associated shapes that have been scanned for naming and build a look-up table. Animates on new blendshape sets, based on their state going high, as well as a latch value being low. Animates last shape left on (found by WhichBlendIsHighDAS) to off state at the same rate, takes the names of the three blend families as input parameters.

WhichBlendIsHighDAS Traverses the three blendshape families and their shapes (number of shapes is dynamic based on number of elements in the blendshape), and reports any weight > 0, (the current shape being used).

ResetBlendsDAS Resets all blend weights to 0 on rewind.

Neutral, Smile, Frown Parse velocity data to determine when to send respective blend "on" signal. Decides which blend to send based on Set Driven Key curve. Will consider a send when latch is low, but waits for an exposed defined delay. If latch goes high from another blend request the request is denied and waits for another latch low.

Depth Parses depth data and keeps the camera dolly the correct distance from the subject avatar.

The process of development of "Face Echoes" has also led to a scoping paper comparing Maya to Unity3D as accessible platforms for developing 3D interactive works, and a set of video tutorials and associated Java and Maya files detailing a workflow and pipeline for using external realtime data transmitted by OSC to Maya. These outcomes were documented and shared on CG Society and YouTube to the Maya community. As Fallman states in "Design-oriented Human-computer Interaction"; "the role of design in HCI must not simply be seen either as a question of problem-solving, as an art-form, or as a bustle with reality: it is on the contrary an unfolding activity which demands deep involvement from the designer" (Fallman, 2003, p. 7).

Discussion

This research studied factors responsible for increased feelings of presence in a facial motion capture animated avatar. Uncanniness is identified as a negative outcome that decreases presence and engagement and care has been taken to avoid this outcome in this research.

During the course of the research there were no studies found that feature a real time facial motion capture animated avatar. The four prototypes detailed in "Phase One: Sketching*"* gave insight into factors that are successful in creating feelings of presence and inviting engagement that can be incorporated into future designs.

The research found that user interactions with virtual props in an environment responsive to facial expressions increases a sense of presence and reduces feelings of uncanniness in users. Steuer predicts this and defines interactivity in the virtual world as "the extent to which users can participate in modifying the form and content of a mediated environment in real time" (Steuer, 1992, p. 12). "Chew" and "Bubble" allow users to directly make new forms and interact with

virtual objects in the virtual world. Users report feelings of presence while interacting with

"Blow", "Chew", and "Bubble", although not in the classic sense of embodiment in the avatar or

presence in the virtual world. Participants reported feeling the virtual props embodied in their

physical body; suggesting that an addition to the phemonological body resulted in increased

presence. This increased engagement through more interaction could be leveraged to better

validate other sensory input channels and increase the overall effectiveness of a simulation.

Combining a facially imitative avatar with an immersive virtual environment using additional

sensory inputs is worthy of further study but was beyond the scope of this research.

It can be argued that there exists a divergent set of contributing attributes for presence in

an environment than for embodiment of a virtual object into the body schema. Body schema

augmentation is at present being explored by Leven et al. in the "Augmented Hand Series";

participants have their hands motion captured and view CG augmentations such as extra fingers

animated in real time on a large screen ("Augmented Hand Series - Interactive Art by Golan

Levin and Collaborators," 2014). The presence that "Chew" and "Bubble" participants reported

suggests that body schema had been altered by the virtual environment. Factors influencing

embodiment of a virtual object into our body schema have not been scoped in existing design

guidelines for virtual environments; but Biocca writes of the effects of changes to the mental

model of self due to alterations in the virtual body that then affect the body schema (Biocca,

1997). Future avatar designers will benefit from continued studies of the effects of the virtual

body on the phenomenal body.

"Flowers" was engaging and elicited an emotional response, but did not embody

participants in the avatar. The simulation did show that the Duchenne smile is a powerful tool

that is able to bridge the uncanny valley. Unlike the results of the study by Tinwell et al. (2011),

the Duchenne smile in the avatar was not found to be uncanny as a result of lack of upper face elements. The Duchenne smile that greets participants in "Face Echoes" is also indicitive of the power of this very human facial expression to positively influence face to face communication and to promote social presence by avoiding feelings of uncanniness.

There were no reports of feelings of presence in participants of "Flowers", even while the virtual environment was being manipulated using facial expressions. The emotional response elicited by the facial feedback may have been overpowering and distracting to the embodiment, or the fact that the virtual environment response was outside the body schema may have been responsible for the lack of comments about presence from the participants. Facial motion capture does offer new methodologies for conducting facial feedback research without the distraction of EMG measurement sensors or the subjectivity of human verification of subject's facial expressions.

Biocca defines social presence as "… when users feel that a form, behavior, or sensory experience indicates the presence of another intelligence" (Biocca, 1997, p. 19). "Face Echoes" creates this response in participants with many waving and saying goodbye after experiencing the simulation and would appear to a have moderate level of social presence; "The amount of social presence is the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another" (Biocca, 1997, p. 19). "Face Echoes" would appear to prove Biocca's statement that "If convincing morphology is present, less intelligence may be required to fool the user into believing that a human intelligence is 'present'" (Biocca, 1997, p. 19).

Biocca (1997) also states that progressive embodiment will depend upon sensory engagement with the number of channels engaged, the fidelity of displays, and saturation and suppression of specific channels being key elements of the engagement; but this study would

argue that higher prioritization and specific application of other elements such as key motor

engagement channels and resolution of body sensors on those channels will better enhance

communication characteristics of CG avatar agents.

Future studies will need to explore face to face avatars from first person point of view in

order to achieve higher levels of communication and the promise of *hyperpresence*, a mode of

communication "where one feels greater 'access to the intelligence, intentions, and sensory

impressions of another' than is possible in the most intimate, face-to-face communication"

(Biocca, 1997, p. 20).

References

Augmented Hand Series - Interactive Art by Golan Levin and Collaborators. (2014). Retrieved

    May 16, 2015, from http://www.flong.com/projects/augmented-hand-series/

AutoEmotive. (n.d.). Retrieved April 2, 2015, from http://autoemotive.media.mit.edu/

Biocca, F. (1997). The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments.

    Journal of Computer-Mediated Communication, *3*(2), 0–0. http://doi.org/10.1111/j.1083-

    6101.1997.tb00070.x

Candy, L. (2006). Practice based research: A guide. *CCS Report*, *1*, 1–19.

Cleland, K. (2006). Media Mirrors and Image Avatars. *Engage: Audience, Art and Interaction*

    (Conference Proceedings) Ernest Edmonds, Lizzie Muller and Deborah Turnbill (eds.),

    University of Technology, Sydney. Retrieved from http://www.kathycleland.com/wp-

    content/uploads/2009/06/cleland_image-avatars-and-media-mirrors.pdf

Cleland, K. (2008). *Image Avatars: Self-other Encounters in a Mediated World*. University of

    Technology, Sydney. Retrieved from http://www.kathycleland.com/wp-

    content/uploads/2009/07/cleland_thesis-2008-image-avatars.pdf

Darwin, C. (1998). *The Expression of the Emotions in Man and Animals*. Oxford University

    Press.

Dimberg, U., & Söderkvist, S. (2011). The Voluntary Facial Action Technique: A Method to

    Test the Facial Feedback Hypothesis. *Journal of Nonverbal Behavior*, *35*(1), 17–33.

    http://doi.org/10.1007/s10919-010-0098-6

Ekman, P. (1992a). An argument for basic emotions. *Cognition & Emotion*, *6*(3-4), 169–

    200. http://doi.org/10.1080/02699939208411068

Ekman, P. (1992b). Facial Expressions of Emotion: New Findings, New Questions.

    Psychological Science, *3*(1), 34–38. http://doi.org/10.1111/j.1467-9280.1992.tb00253.x

Ekman, P., & Friesen, W. V. (1977). Facial action coding system.

Ekman, P., & Friesen, W. V. (2003). *Unmasking the Face: A Guide to Recognizing Emotions*

    from Facial Clues. ISHK.

Ekman, P., Friesen, W. V., & Levenson, R. W. (1990). Voluntary Facial Action Generates

    Emotion-Specific Autonomic Nervous System Activity. *Psychophysiology*, *27*(4), 363–

    384. http://doi.org/10.1111/j.1469-8986.1990.tb02330.x

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity

    distinguishes among emotions. *Science*, *221*(4616), 1208–1210.

    http://doi.org/10.1126/science.6612338

Emotient - Facial Expression Recognition Software. (n.d.). Retrieved June 9, 2014, from

    http://www.emotient.com/

Faceshift. (n.d.). Retrieved June 9, 2014, from http://www.faceshift.com/

Facial Expression Analysis Recognition-Affectiva. (n.d.). Retrieved June 9, 2014, from

    http://www.affdex.com/

Fallman, D. (2003). Design-oriented Human-computer Interaction. In *Proceedings of the*

    SIGCHI Conference on Human Factors in Computing Systems (pp. 225–232). New York,

    NY, USA: ACM. http://doi.org/10.1145/642611.642652

Freud, S., Strachey, J., Cixous, H., & Dennomé, R. (1976). Fiction and Its Phantoms: A Reading

    of Freud's Das Unheimliche (The "Uncanny"). *New Literary History*, *7*(3), 525–645.

    http://doi.org/10.2307/468561

Gilbert, J. (n.d.). A Running List of Places That Have Banned Selfie Sticks. Retrieved May 5,

  2015, from https://www.yahoo.com/tech/a-running-list-of-places-that-have-banned-

  selfie-113345971759.html

GTA 5 Selfies. (n.d.). Retrieved May 15, 2015, from http://gta5selfies.com/

Intel® RealSense[TM] Technology. (n.d.). Retrieved November 8, 2014, from

  http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-

  overview.html

James Cameron's Avatar. (n.d.). Retrieved May 5, 2015, from http://origin.avatarmovie.com/

Jentsch, E. (1997). On the psychology of the uncanny (1906) 1. *Angelaki: Journal of the*

  Theoretical Humanities, *2*(1), 7–16.

Kim, T., & Biocca, F. (1997). Telepresence via Television: Two Dimensions of Telepresence

  May Have Different Connections to Memory and Persuasion.[1]. *Journal of Computer-*

  Mediated Communication, *3*(2), 0–0. http://doi.org/10.1111/j.1083-6101.1997.tb00073.x

Kinect for Windows. (n.d.). Retrieved May 10, 2015, from https://www.microsoft.com/en-

  us/kinectforwindows/

Larman, C., & Basili, V. R. (2003). Iterative and Incremental Developments: A Brief History.

  Computer, *36*(6), 47–56. http://doi.org/10.1109/MC.2003.1204375

Li, H., Weise, T., & Pauly, M. (2010). Example-based Facial Rigging. In *ACM SIGGRAPH 2010*

  Papers (pp. 32:1–32:6). New York, NY, USA: ACM.

  http://doi.org/10.1145/1833349.1778769

Maddock, Steve, Edge, James, & Sanchez, Manuel. (2000). *Movement Realism in Computer*

  Facial Animation. Cambridge, Mass.: MIT Press. Retrieved from

http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&A
N=32650

Massive Software – Simulating Life. (n.d.). Retrieved May 10, 2015, from
http://www.massivesoftware.com/

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE
Robotics Automation Magazine*, *19*(2), 98–100.
http://doi.org/10.1109/MRA.2012.2192811

Pandzic, I. S., & Forchheimer, R. (2003). *MPEG-4 Facial Animation: The Standard,
Implementation and Applications*. John Wiley & Sons.

Perry, T. S. (2014). Leaving the uncanny valley behind. *Spectrum, IEEE*, *51*(6), 48–53.

Pollick, F. E. (2010). In Search of the Uncanny Valley. In P. Daras & O. M. Ibarra (Eds.), *User
Centric Media* (pp. 69–78). Springer Berlin Heidelberg. Retrieved from
http://link.springer.com.ezproxy.aut.ac.nz/chapter/10.1007/978-3-642-12630-7_8

Processing.org. (n.d.). Retrieved May 9, 2015, from https://processing.org/

Sagar, M. (2006). Facial Performance Capture and Expressive Translation for King Kong. In
ACM SIGGRAPH 2006 Sketches. New York, NY, USA: ACM.
http://doi.org/10.1145/1179849.1179882

Saltz, J. (2014). Art at Arm's Length: A History of the Selfie. *New York Magazine*, *47*(2), 71–75.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (Vol. 5126).
Basic books.

Schultze, U. (2010). Embodiment and presence in virtual worlds: a review. *Journal of
Information Technology*, *25*(4), 434–449. http://doi.org/10.1057/jit.2010.25

Soussignan, R. (2002). Duchenne Smile, Emotional Experience, and Autonomic Reactivity: A

    Test of the Facial Feedback Hypothesis. *Emotion March 2002*, *2*(1), 52–74.

    http://doi.org/10.1037/1528-3542.2.1.52

Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of*

    Communication, *42*(4), 73–93. http://doi.org/10.1111/j.1460-2466.1992.tb00812.x

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the

    human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of*

    Personality and Social Psychology, *54*(5), 768–777. http://doi.org/10.1037/0022-

    3514.54.5.768

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion

    and perception of the Uncanny Valley in virtual characters. *Computers in Human*

    Behavior, *27*(2), 741–749. http://doi.org/10.1016/j.chb.2010.10.018

Toolkit for Sensing People in Spaces. (n.d.). Retrieved May 9, 2015, from http://www.tsps.cc/

Weise, T., Bouaziz, S., Li, H., & Pauly, M. (2011). Realtime Performance-based Facial

    Animation. In *ACM SIGGRAPH 2011 Papers* (pp. 77:1–77:10). New York, NY, USA:

    ACM. http://doi.org/10.1145/1964921.1964972

Wortham, J. (2013, October 19). My Selfie, Myself. *The New York Times*. Retrieved from

    http://www.nytimes.com/2013/10/20/sunday-review/my-selfie-myself.html

Appendix A. Video Consent Forms

Consent and Release Form

For use with photographic projects

AUT
UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project title:	***Avatar Mirrors***
Project Supervisor:	***James Charlton***
Researcher:	***Donald Smith***

○ I have read and understood the information provided about this research project in the Information Sheet dated 15 01 2015.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that I may withdraw myself, my image, or any other information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ If I withdraw, I understand that all relevant information will be destroyed.

○ I permit the artist to use the photographs and video that are part of this project and any other reproductions, adaptation or 3D data derived from them, either complete or in part, alone or in conjunction with any wording and/or drawings solely and exclusively for (a) the artist's portfolio; and (b) educational exhibition and examination purposes and related design works.

○ I understand that the photographs, video and 3D data will be used for academic purposes only and will not be published in any form outside of this project without my written permission.

○ I understand that any copyright material created by the photographic sessions is deemed to be owned by the artist and that I do not own copyright of any of the photographs, video or 3D data.

○ I agree to take part in this research.

Participant's signature: ....

Participant's name: .... Jenna Gavin ....

Participant's Contact Details (if appropriate):

……………………………………………………………………………..

……………………………………………………………………………..

……………………………………………………………………………..

……………………………………………………………………………..

Date: 25 May 2015

Note: The Participant should retain a copy of this form.

This version was last edited on 8 November 2013

21 May 2015

Consent and Release Form

For use with photographic projects

A U T
U N I V E R S I T Y
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project title:	***Avatar Mirrors***
Project Supervisor:	***James Charlton***
Researcher:	***Donald Smith***

O   I have read and understood the information provided about this research project in the Information Sheet dated 15 01 2015.

O   I have had an opportunity to ask questions and to have them answered.

O   I understand that I may withdraw myself, my image, or any other information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

O   If I withdraw, I understand that all relevant information will be destroyed.

O   I permit the artist to use the photographs and video that are part of this project and any other reproductions, adaptation or 3D data derived from them, either complete or in part, alone or in conjunction with any wording and/or drawings solely and exclusively for (a) the artist's portfolio; and (b) educational exhibition and examination purposes and related design works.

O   I understand that the photographs, video and 3D data will be used for academic purposes only and will not be published in any form outside of this project without my written permission.

O   I understand that any copyright material created by the photographic sessions is deemed to be owned by the artist and that I do not own copyright of any of the photographs, video or 3D data.

O   I agree to take part in this research.

Participant's signature: ....……………………………………………………

Participant's name: Phil James   ....……………………………………………………

Participant's Contact Details (if appropriate):

philljames@gmail.com
……………………………………………………………………………………………..

+6421708684
……………………………………………………………………………………………..

……………………………………………………………………………………………..

……………………………………………………………………………………………..

Date: 26 May 2015

Note: The Participant should retain a copy of this form.

This version was last edited on 8 November 2013

21 May 2015

Consent and Release Form

For use with photographic projects

AUT
UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project title: **Avatar Mirrors**

Project Supervisor: **James Charlton**

Researcher: **Donald Smith**

○ I have read and understood the information provided about this research project in the Information Sheet dated 15 01 2015.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that I may withdraw myself, my image, or any other information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ If I withdraw, I understand that all relevant information will be destroyed.

○ I permit the artist to use the photographs and video that are part of this project and any other reproductions, adaptation or 3D data derived from them, either complete or in part, alone or in conjunction with any wording and/or drawings solely and exclusively for (a) the artist's portfolio; and (b) educational exhibition and examination purposes and related design works.

○ I understand that the photographs, video and 3D data will be used for academic purposes only and will not be published in any form outside of this project without my written permission.

○ I understand that any copyright material created by the photographic sessions is deemed to be owned by the artist and that I do not own copyright of any of the photographs, video or 3D data.

○ I agree to take part in this research.

Participant's signature: …………………………………………………………………………………………

Participant's name: …………… EMILE DRESCHER …………………………………………………………

Participant's Contact Details (if appropriate):

…………………………………………………………………………………………..

…………………………………………………………………………………………..

…………………………………………………………………………………………..

…………………………………………………………………………………………..

Date: **2/06/2015**

Note: The Participant should retain a copy of this form.

This version was last edited on 8 November 2013

21 May 2015

AUT
U N I V E R S I T Y
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Consent and Release Form

For use with photographic projects

Project title:	***Avatar Mirrors***
Project Supervisor:	***James Charlton***
Researcher:	***Donald Smith***

○ I have read and understood the information provided about this research project in the Information Sheet dated 15 01 2015.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that I may withdraw myself, my image, or any other information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ If I withdraw, I understand that all relevant information will be destroyed.

○ I permit the artist to use the photographs and video that are part of this project and any other reproductions, adaptation or 3D data derived from them, either complete or in part, alone or in conjunction with any wording and/or drawings solely and exclusively for (a) the artist's portfolio; and (b) educational exhibition and examination purposes and related design works.

○ I understand that the photographs, video and 3D data will be used for academic purposes only and will not be published in any form outside of this project without my written permission.

○ I understand that any copyright material created by the photographic sessions is deemed to be owned by the artist and that I do not own copyright of any of the photographs, video or 3D data.

○ I agree to take part in this research.

Participant's signature: ....

Participant's name: **Hengbo Wang**

Participant's Contact Details (if appropriate):

Email: e@edisonw.com

Mobile: 027-9301903

....

Date: **2/5/2015**

Note: The Participant should retain a copy of this form.

21 May 2015

AUT
UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Consent and Release Form

For use with photographic projects

Project title:	**Avatar Mirrors**
Project Supervisor:	**James Charlton**
Researcher:	**Donald Smith**

○   I have read and understood the information provided about this research project in the Information Sheet dated 15 01 2015.

○   I have had an opportunity to ask questions and to have them answered.

○   I understand that I may withdraw myself, my image, or any other information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○   If I withdraw, I understand that all relevant information will be destroyed.

○   I permit the artist to use the photographs and video that are part of this project and any other reproductions, adaptation or 3D data derived from them, either complete or in part, alone or in conjunction with any wording and/or drawings solely and exclusively for (a) the artist's portfolio; and (b) educational exhibition and examination purposes and related design works.

○   I understand that the photographs, video and 3D data will be used for academic purposes only and will not be published in any form outside of this project without my written permission.

○   I understand that any copyright material created by the photographic sessions is deemed to be owned by the artist and that I do not own copyright of any of the photographs, video or 3D data.

○   I agree to take part in this research.

Participant's signature: June Kim   ....

Participant's name: June Kim   ....

Participant's Contact Details (if appropriate):

301-88 The Strand, Parnell, Auckland
....

....

....

....

Date:

Note: The Participant should retain a copy of this form.