Sabrina Bleecker Caldwell
The Australian National University
Research School of Computer Science

November, 2016

Framing digital image credibility: image manipulation problems, perceptions and solutions

A thesis submitted for the degree of Doctor of Philosophy of
The Australian National University

I certify that this thesis is my own original work, resulting from research undertaken by me as a PhD candidate at the Australian National University

Words: 47250

Sabrina Caldwell

Sabrina Bleecker Caldwell

Acknowledgements

To create an active tempo in my writing and to reflect that it is my own research, I have written this thesis largely in first person, but I wish to acknowledge that just as it takes a village to raise a child, it takes a community to provide the intellectual contributions, financial and logistical support, and pastoral care required for the epic journey that is a PhD.

I am grateful to many people for their thoughtful and inspirational assistance to me in undertaking this research. First and foremost I wish to acknowledge my principal supervisor and advisor Professor Tamás Gedeon for his sage advice and practical support. Secondly, I want to express my appreciation to Dr Richard Jones, whose support was also invaluable and treasured. Together, Tom and Richard provided me with years of encouragement, consultation, understanding, and even friendship, and provided me with many exciting opportunities to advance and promote my research. I also thank Dr Martyn Jolly of the School of Art for his expertise in photography and cross-disciplinary perspectives in respect of my investigations.

As my research progressed, I found myself working with a fellow PhD student, Leana Copeland, who also used eye gaze as a tool of investigation (in her case for reading comprehension), and together we were able to undertake joint experimentation in which she gained outcomes for her research and I gained outcomes for mine. It was wonderful working with her and I hope that an opportunity arises to do so again in future.

On a personal note, I wish to thank my husband Brian Caldwell for his endless and steadfast support. It is seldom that a man is called upon to endure the trials of supporting a wife undertaking PhD studies not just once, but twice.

Lastly, I want to thank the Research School of Computer Science, the College of Engineering and Computer Science and ANU for their faith in me, and for allowing me this extraordinary opportunity to follow my passion to find a practical solution to the problem of credibility in photography, and the visual information it captures, for posterity.

Abstract

Image manipulation is subverting the credibility of photographs as a whole. Currently there is no practical solution for asserting the authenticity of a photograph. People express their concern about this when asked but continue to operate in a 'business as usual' fashion.

While a range of digital forensic technologies has been developed to address falsification of digital photographs, such technologies begin with 'sourceless' images and conclude with results in equivocal terms of probability, while not addressing the meaning and content contained within the image.

It is interesting that there is extensive research into computer-based image forgery detection, but very little research into how we as humans perceive, or fail to perceive, these forgeries when we view them. The survey, eye-gaze tracking experiments and neural network analysis undertaken in this research contribute to this limited pool of knowledge.

The research described in this thesis investigates human perceptions of images that are manipulated and, by comparison, images that are not manipulated. The data collected, and their analyses, demonstrate that humans are poor at identifying that an image has been manipulated. I consider some of the implications of digital image manipulation, explore current approaches to image credibility, and present a potential digital image authentication framework that uses technology and tools that exploit social factors such as reputation and trust to create a framework for technologically packaging/wrapping images with social assertions of authenticity, and surfaced metadata information.

The thesis is organised into 6 chapters.

Chapter 1: Introduction

I briefly introduce the history of photography, highlighting its importance as reportage, and discuss how it has changed from its introduction in the early 19th century to today. I discuss photo manipulation and consider how it has changed along with photography. I

describe the relevant literature on the subject of image authentication and the use of eye gaze tracking and neural nets in identifying the role of human vision in image manipulation detection, and I describe my area of research within this context.

Chapter 2: Literature review

I describe the various types of image manipulation, giving examples, and then canvas the literature to describe the landscape of image manipulation problems and extant solutions, namely:

- the nature of image manipulation,
- investigations of human perceptions of image manipulation,
- eye gaze tracking and manipulated images,
- known efforts to create solutions to the problem of preserving unadulterated photographic representations and the meanings they hold.

Finally, I position my research activities within the context of the literature.

Chapter 3: The research

I describe the survey and experiments I undertook to investigate attitudes toward image manipulation, research human perceptions of manipulated and unmanipulated images, and to trial elements of a new wrapper-style file format that I call .msci (mobile self-contained image), designed to address image authenticity issues.

Methods, results and discussion for each element are presented in both explanatory text and by presentation of papers resulting from the experiments.

Chapter 4: Analysis of eye gaze data using classification neural networks

I describe pattern classifying neural network analysis applied to selected data obtained from the experiments and the insights this analysis provided into the opaque realm of cognitive perception as seen through the lens of eye gaze.

Chapter 5: Discussion

I synthesise and discuss the outcomes of the survey and experiments.

I discuss the outcomes of this research, and consider the need for a distinction between photographs and photo art. I offer a theoretical formula within which the overall authenticity of an image can be assessed. In addition I present a potential image authentication framework built around the .msci file format, designed in consideration of my investigation of the requirements of the image manipulation problem space and the experimental work undertaken in this research.

Chapter 6: Conclusions and future work

This thesis concludes with a summary of the outcomes of my research, and I consider the need for future experimentation to expand on the insights gained to date. I also note some ways forward to develop an image authentication framework to address the ongoing problem of image authenticity.

Table of Contents

Table of Figures

Tables

PhD Research Papers, Presentations, Awards

Papers – first author

2016	**Caldwell, Sabrina** and Tamás Gedeon. (2016) Boldly going where no higher educators have gone before: a review of the 1st higher education advances conference, Valencia Spain. Procedia Volume 228 pp 348-355
2015	**Caldwell, Sabrina**, Tamás Gedeon, Richard Jones, Leana Copeland. (2015) Imperfect understandings: a grounded theory and eye gaze investigation of human perceptions of manipulated and unmanipulated digital images. Conference Proceedings of 3rd International Conference on Human-Computer Interaction, Barcelona, Spain July 13-14, 2015 (**Best Paper Award**). http://avestia.com/EECSS2015_Proceedings/files/papers/MHCI308.pdf.
2015	**Caldwell, S.B** and T.D. Gedeon. Optimising Peer Marking with Explicit Training: from Superficial to Deep Learning. 1st International Conference on Higher Education Advances. DOI: http://dx.doi.org/10.4995/HEAd15.2015.441 pp 626-631.
2014	**Sabrina Caldwell**, Tom Gedeon, Richard Jones, Martin Henschke. Comparing eye gaze tracking to reported perceptions of manipulated and unmanipulated digital images. Australian Journal of Intelligent Information Processing Systems. Vol 14 No 3. pp 26-36.
2012	**Sabrina Caldwell** (2012) Tinkering with the Truth: Investigations into the Implications of Digital Image Manipulation and the Technologies of Image Credibility. Poster. Digital Humanities Australasia conference.

Papers – contributing author

2016	Leana Copeland, Tamás Gedeon and **Sabrina Caldwell**. (2016) Mitigating distractions during online reading: an explorative study. IEEE International Conference on Systems, Man and Cybernetics (SMC)
2015	Copeland, Leana, Tom Gedeon and **Sabrina Caldwell**. Effects of Text Difficulty on Predicting Reading Comprehension from Eye Movements. Proceedings of 6th IEEE Conference on Cognitive Infocommunications. Győr, Hungary. October 19-21, 2015.
2014	Copeland, L, Gedeon, TD, **Caldwell, S** (2014) Framework for Dynamic Text Presentation in eLearning, 6th International conference on Intelligent Human Computer Interaction, IHCI 2014, 4 pages.
2013	Martin Henschke, Tom Gedeon, Richard Jones, **Sabrina Caldwell** and Dingyun Zhu. (2013) Wands are Magic: A comparison of devices used in 3D pointing interfaces. Technical Report TR-CS-13-01, Research School of Computer Science, College of Engineering and Computer Science, The Australian National University, 6 pages.

Presentations

2016	Framing digital image credibility: image manipulation problems, perceptions and solutions. ANU Research School of Computer Science. November
2015	Global trends in higher education advances: news from the 1st International Conference on Higher Education Advances 24-26 June 2015, Valencia, Spain" ANU Student Administration and Services (SAS) Forum, September
2014, 2012	Issues of Digital Image Credibility and Authentication. HCC and Friends, ANU Research School of Computer Science. June
2013	Using Wattle with Excel to gain efficiency and consistency in marking student websites Moodleposium, September
2012	Digital Image Credibility. Canon Information Systems Research Australia (CISRA),
2012	Digital Image Credibility and Authentication, School of Art, ANU College of Arts and Social Sciences

Awards

2015	Best paper for Imperfect understandings: a grounded theory and eye gaze investigation of human perceptions of manipulated and unmanipulated digital images, 3rd International Conference on Human-Computer Interaction
2014	Nominated for Teaching Excellence Award for Contributions to Programs that Enhance Learning
2013	Certificate, Foundations of University Teaching and Learning
2013	Vice Chancellor's Teaching Enhancement Grant
2012	Certificate with merit, Graduate Teaching Program

Preamble

A confusing point of terminology in any discussion of photographs and images is the distinction between these terms. The Oxford English Dictionary definition shows the noun *image* to be highly overloaded, with definitions aggregated over lengthy time reaching back at least to the Anglo-Norman and Old French word *himage*. Amongst the 26 definitions and sub-definitions, 3c is the most representative of photography:

3.c *A physical or digital representation of something, originally captured using a camera from visible light, and typically reproduced on paper; displayed on a screen, or stored as a computer file. More generally: any picture or graphic (regardless of origin) displayed on a computer monitor, television, etc., or reproduced in printed form.*

By contrast, the noun *photograph* has just 2 definitions; definition 2 is only figurative, therefore the official OED definition of a photograph is:

1. *A picture or image obtained by photography; (originally) a picture made using a camera in which an image is focused on to sensitive material and then made visible and permanent by chemical treatment; (later also) a picture made by focusing an image and then storing it digitally.*

While there are deficiencies in this definition (not all chemically produced photographs require a camera, and the phrase ‘focusing an image’ is rather vague), the intent appears to be that a photograph is an image that captures light reflected from our physical world. This intent is also evident in the subset definition 3.c above of the term *image*. From this I surmise that all *photographs* are *images*, but not all *images* are *photographs*.

Many of the definitions of *image* are not relevant to my research, which focusses on photographs, manipulated images, and photo art. Accordingly, for the purposes of this thesis, the term *image* refers to the set of photograph, manipulated image and photo art. The term *photograph* refers to the type of image that is created (either chemically or digitally) by capturing light reflected from the physical world around us.

Another term frequently used (perhaps coined) in this thesis is the word unmanipulated, to denote a photograph that has been left in its original state. This fills a void in expressing the distinction between the two types of photograph (unmanipulated and manipulated) and is in keeping with the OED statement that the prefix un- is “very extensively employed in English, as in the other Germanic languages, and is now the one which can be used with the greatest freedom in new formations”. (OED, 2016)

1 Introduction

Understanding the extent to which an image can be believed is becoming increasingly important as ‘enhanced’ digital images proliferate, and purport to represent the reality of individuals and our environments, record significant human events, and are used as supportive evidence in areas as diverse as academic research and law enforcement. Issues with image credibility range from confusing modifications of everyday digital images, through to misrepresentations of war zones. At the same time, there are as yet poorly understood shifts in the roles images play in human endeavour, as more and more photographs are either ‘born’ digital or are converted to digital.

Stories involving high profile misleading images appear again and again in the media. The modified Beirut photograph (Figure 1-1) in which an over-zealous photographer used photo editing to expand a warzone is merely one example in a long chronology of manipulated military propaganda and quasi-reportage that reaches back to Matthew Brady’s photographs of repositioned American Civil War dead, and forward to today with North Korea’s frequent use of image ‘cloning’ to exaggerate the size of its military force.

“Smoke billows from burning buildings destroyed during an overnight Israeli air raid on Beirut’s suburbs August 5, 2006. Many buildings were flattened during the attack.” REUTERS/Adnan Hajj

Figure 1-1: Cloned smoke columns, Beirut

Modified photographs appear in science: from NASA published images of the Mars landscape as red instead of its true brown colour to meet public expectations, through to spliced images passed as evidence in medical research. It is common in politics, from Stalin’s erasures of political opponents in his photographs to ‘doctored’ salacious images of Australia’s One Nation senate candidate Pauline Hanson appearing in Australian newspapers.

It is clear that manipulated images are not new, but what *is* new is the rate and volume at which they are being produced, and the ease of international distribution afforded them by the internet.

Problems with digital images are being recognized in varying disciplines such as law enforcement and the media, but these problems do not as yet seem to be connected in public perception as an overall issue with digital photography. Instead, these issues seem to be treated as isolated instances.

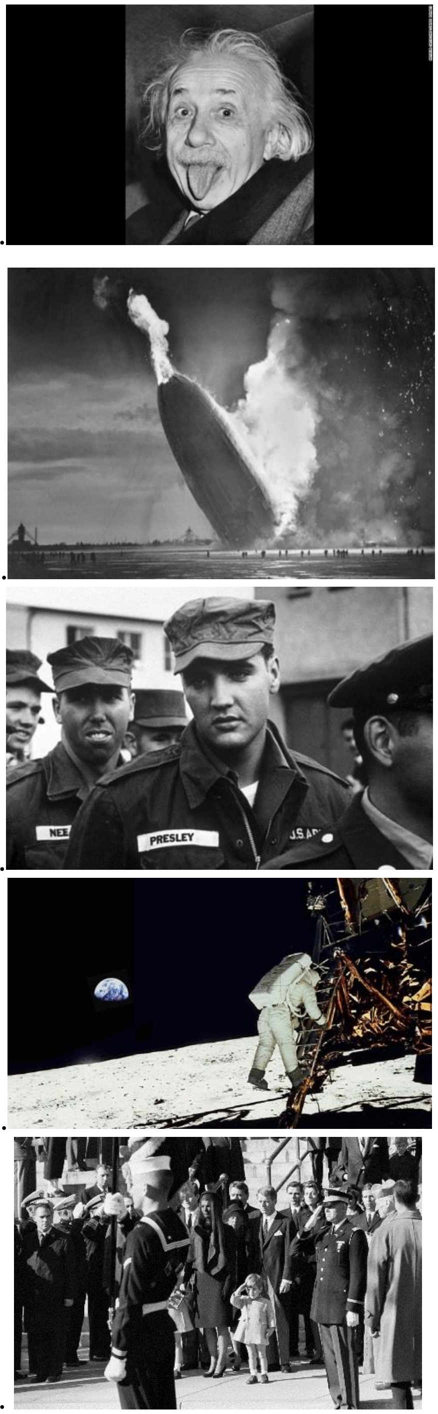
When challenged with issues of misrepresentation in photography, photographers respond that photographs 'have always been subjective' and point to iconic photographic hoaxes such as the infamous Loch Ness plastic toy monster (Siegel, 1994) or the cardboard Cottingley fairies (Banks, 2013).

In day-to-day personal photo production, new ways of altering photos are transforming the ubiquitous 'snapshot.' Instagram provides, even encourages 'on-the-fly' photo effects. Photographs can be tremendously altered before they even leave the camera. Many cameras provide 'in camera' options to create images pre-altered by filters. One new software system bundled in some cameras, PhotoDirector, post processes multiple photographs to create perfect group photos by swapping faces to ensure the maximum number of smiling faces, enhances skin colour by lightening or tanning, and stiches seamless panoramas from separate photos (Cyberlink, 2016). This means that in some cases modern photographs are not just 'born digital,' they are 'born digitally enhanced.'

It is easy to see that issues of image manipulation are complex, comprising an amalgam of factors ranging from technological to monetary, contextual to cultural, and historic to future-facing. These conceptual and technical challenges appear overwhelming, rooted in photographic practices stretching back to the inception of photography in 1839, and racing into the future on successive waves of technological progress in pre- and post-image processing. At the moment, it even seems to some that the problem has bolted away from our ability to manage it, while others query, "what problem?"

However, it is possible that the advent of the self-same technologies of ubiquitous computing and the internet that have given rise to escalating issues of digital image credibility also hold the key to their solution. It may be that we have the ability to restore the credibility of photographs in future by using computing technologies to capture an unchanging version of the original photograph, like a fly in amber, along with the desired presentation image and provenance properly ascribed to it.

The following research spans the boundaries of computer science and the social sciences. Its cross-disciplinary approach provides a holistic perspective on a subject, with a net cast necessarily wide to accommodate the many and varied facets of image tampering, its causes, effects, and redress.



a.
b.
c.
d.
e.

Figure 1-2: Historic images a) Albert Einstein b) Hindenburg disaster c) Elvis Presley enlists d) Aldrin moonwalk e) JFK's son salutes

1.1 Why photographs are important

It is beyond the scope of this thesis to exhaustively catalogue all the roles played by representative photography in our society. I claim that it is only necessary to demonstrate that representative photographs *do* have important roles to play. I will do this by briefly mentioning three of the key areas in which photographs are indispensable: 1) history/memories and the human experience, 2) evidence and justice, and 3) science and medicine.

1.1.1 History, memories and the human experience

Photographs are used extensively in preserving knowledge of history. To intuitively understand this, one has only to consider the photographs capturing the explosion of the Hindenburg zeppelin in 1937 (Taylor, 2012), the first moonwalk (NASA, 1969), and photos of historic figures like Albert Einstein (Sasse,

1951) and Elvis Presley (Lebeck, 1958), as well as John F. Kennedy Jr saluting his father at JFK's state funeral (Stearnes, 1963) (Figure 1-2).

Photographs memorialise human history, and also personal memories. When the frequent bushfires of Australia destroy homes, it is often lost photographs that are the first non-living things that survivors mourn (Harrison, 2002). Photos are an aid to our memories, a medium of social relationships, and an important conduit of self-expression and self-presentation (House, Davis, Takhteyev, Ames, & Finn, 2004). Be they casual snapshots or carefully composed photographs, without photographs to capture our lives' memorable moments, our experience of our personal histories would be much poorer.

1.1.2 Evidence, security and justice

> *We cannot conceive of a more impartial and truthful witness than the sun, as its light stamps and seals the similitude of the wound on the photograph put before the jury; it would be more accurate than the memory of witnesses, and as the object of all evidence is to show truth, why should not this dumb witness show it?*
>
> - Franklin v. State of Georgia, 69 Ga. 36; 1882 Ga.
> - As quoted in *Photographs as Evidence* (Meskin & Cohen, 2007)

During most of the 20th century photographic evidence has played a pivotal role in the lives of anyone touched by criminal or justice issues. Arrested criminals have been photographed for future reference since the 1840s, starting only a few years after the invention of photography.

Figure 1-3: 9/11 terrorist attack on World Trade Center

In the 21st century, the use of security cameras, and surveillance images in public security has grown significantly, particularly in the wake of 9/11 and other terrorist attacks around the world.

Photographs are accepted as faithful reportage supporting law suits, prosecution of criminals, and exposing the malfeasance of government officials and corporations. They have also been key to court proceedings since the late 19th century; often providing the pivotal evidence that determines the fate of defendants.

They may be presented at court as witness supported images, or they may be standalone images that speak on their own behalf; the authenticity and credibility of such standalone digital images introduced in court cases are particularly sensitive in trials, and vulnerable to challenge by defendants.

1.1.3 Science and medicine

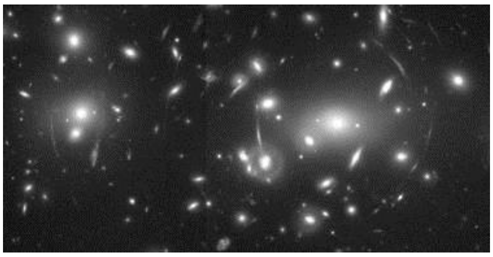


Figure 1-4: Abell 2218, Hubble telescope (Couch, 1995)

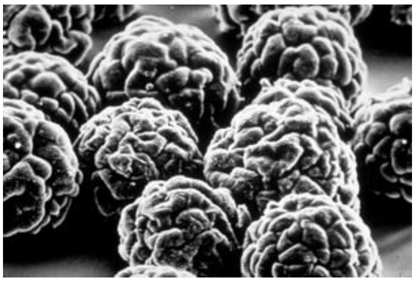


Figure 1-5: Wild poliovirus photographed through microscope (Global Polio Eradication Initiative, 2010)

Photographs provide us with evidence about visual properties of the world around us, and by that means we gain knowledge about the physical world.

We launch telescopic cameras like the Hubble into space to photograph distant galaxies and indirectly understand time. We amass image repositories of microorganisms and viruses to better understand how to keep ourselves healthy.

Photographic images are used in scientific and medical research and are published as factual data supporting lines of investigation in journals and distributed to the public for general interest and news.

1.1.4 Visual culture

Digital images are being used in place of text to communicate. It is common for instance for someone to use their Smartphone to snap a photo of their latte and perhaps their mother as well and send it to their friend on the spot, rather than to later write a note or letter to a friend or relative that they had a nice coffee visit with their mum. Each one of these ‘digital traces’ is a potential memory (Hand, 2014).

1.2 Photo manipulation: a two-edged sword

It is important to acknowledge the positive benefit of image manipulation in many fields such as enhancing images to interrogate visual data to a greater extent than possible with the natural eye. In that fashion we can exploit new opportunities for discovery.

Through enhancement of images we can often see better and differently into some of the important images we capture. In fact, some frequencies of electromagnetic radiation are outside the range of human perception, and *require* false colour representation to be visible to us. In other cases, simply looking at an image after it has been post processed provides new understanding and other knowledge benefits.

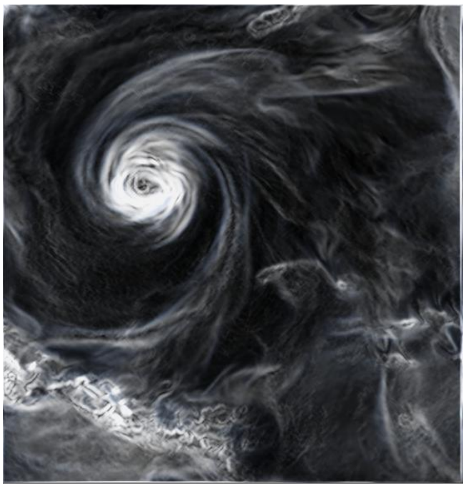


Figure 1-6: Original image of Hurricane Katrina, rendered using silhouette enhancement to better outline key features

As an example of the beneficial use of image enhancement, Figure 1-6 shows an enhanced image of 2005's devastating Hurricane Katrina. The image was generated with a silhouette enhancement technique to preserve the key features of the hurricane – eye, eyewall and rainbands – while removing obscuring features that may make it difficult to predict the hurricane's impact and direction (Joshi, Caban, Rheingans, & Sparling, 2009). Image manipulations like these can have positive benefits and can even save lives.

Figure 1-7: 1989 Oprah Winfrey cover of TV Guide

But there are also many issues arising in respect of our ability to easily alter our images. Consider the unauthorized splicing of the head of Oprah Winfrey, one of the most powerful women in the US media, onto the body of white actress Ann-Margret (Figure 1-7). There are multiple ethical and cultural issues raised by such a manipulation – perhaps most significantly this spliced image subverts Winfrey's status as a high-achieving black woman role model. Further, the image reinforces body image stereotypes (Winfrey's normal body-type is not similar to Ann-Margret's idealized figure), and by also splicing a pile of cash into the montage, seems to give evidence that Winfrey possesses large quantities of cash and is willing to display herself with it; something that is at odds with her public image as a caring champion of society's disadvantaged and disenfranchised.

In another example, representations of O.J. Simpson by competing news magazines *Newsweek* and *Time* created a national stir in 1994 in the United States when *Time* chose to present a menacing looking Simpson on their cover, with the minatory ambience created by darkening his colour.

Figure 1-8: O.J. Simpson mugshot, cover of TIME magazine, cover of Newsweek magazine 1994

This was an extraordinary moment in journalism in which a high-profile news magazine a) blatantly played on racism by using dark skin colour as a metaphor for criminality, and b) manipulated that colour to visually imply a high degree of guilt of a suspect prior to his trial on the front cover of their magazine. These images were firmly in the public eye even before a jury was empanelled for his trial. Simpson was later found not guilty in criminal court of the murders of Nicole Brown Simpson and Ronald Goldman, though he was found guilty in a civil court.

The impacts of manipulated images can extend into altered perceptions of reality. Tampered images presented to people of purported past events in their lives or of public events have been shown to change the way people remember events, in short, they implant false memories. Frenda et al led their study subjects to ‘recall’ having seen Obama shaking hands with former Iranian President Mahmoud Ahmadinejad after being shown a doctored photo of this event purportedly happening. (Frenda, Knowles, Saletan, & Loftus, 2013). Wade et al led people to ‘remember’ a fictitious balloon ride from their childhood, a ride they never took (Garry & Wade, 2005).

The use of potentially manipulated images in law and justice has also been controversial, especially when defendants can question the credibility of the images used in their prosecution. There have also been questions about the new use of images in judges’ case summaries, which were traditionally *written*, with concerns that issues of image manipulation or simply visual misinterpretations can arise (Dellinger, 1997).

There is a great deal of concern in the medical community about the potential for fraud and harm. In considering the benefits and detriments of image manipulation, Suvarna and Ansary note a significant positive that, with proper enhancement, analysis of histopathological images led to better health outcomes, and that “…the false image is of greater value in terms of patient management, although the original image recorded is also part of the analysis overall.” They equally point to the fraud perpetrated on the scientific community by the scientists Brach and Hermann, who deliberately falsified at least 4 publicly funded cancer studies by splicing together different autoradiographs (Steimle, 1998; Suvarna & Ansary, 2001). Dermatologists Cutrone and Gimalt expressed concern that people in need of health care may be lured to purchase dangerous and expensive medications by online suppliers posting fraudulent images portraying seeming miraculous cures of diseases like dermatitis, vitiligo, and psoriasis (Cutrone & Grimalt, 2001). Patients themselves can manipulate images for medical insurance fraud (Anoop, 2015).

It seems we don’t always equate a digital image with the *data* it actually records. “Scientists must remember that digital images are numerically sampled data that represent the state of a specific sample when examined with a specific instrument” (Cromey, 2013).

Misleading others using photo manipulation can be an issue even at the highest levels of government. In 2014, the Malaysian government released images of suspects in the MH370 passenger plane bombing in which the legs of the suspects are clearly and inexplicably identical. (Pollard, 2014). The authorities were caught out giving the false information when a reporter noted that photos of two suspects from the MH370 airline disaster had the same exact green clad legs. This is particularly interesting since the shadow appearing on the left version of the legs would have to have been removed for the right version.

Figure 1-9 - MH370 suspects have suspect legs (scan of newspaper photos)

North Korea regularly issues photographs in which the might of its military is enhanced by cloning warships or improving the performance of missile launches. One recent propagandist photograph was released in May 2015 as misleading ‘proof’ of an underwater test-fire of a missile launch from a submarine. Aerospace engineers from Schmucker Technologie in Germany pointed out the that reflections of the missile exhaust flame were not aligned with the missile, and suggested that the missile had been launched from a submerged barge (Culbertson, 2015).

It is not just these high profile image manipulations that are of concern. Ubiquitous image editing software makes it easy for anyone to modify a digital photo, and these types of manipulations take place in thousands of households. Snap photographers might remove a wrinkle from a loved one’s face, erase the unsightly trash can lurking behind a field of tulips, or crop out a neighbour who has infiltrated a family photo. Or they might go further: change a grey sky to blue, insinuate a friend into the arms of her favourite celebrity, or layer several photographs together to create images ranging from indiscernibly similar to a representative photo through to an obvious ‘mashup.’

Minor or major, these changes reduce an image’s credibility as a report of real people and events. These increasingly uncertain images are collectively telling the story of our personal histories and constructing our broader social and historical narratives. The tools available that encourage this photo-play are insidious. An Instagram-filtered image (Figure 1-10) is not really a photograph, although it may be a fun piece of photo art in its own right. But if in later years we wish to remember a day, a person, or an event as it actually was, Instagram filtered images alone will not achieve this for us.

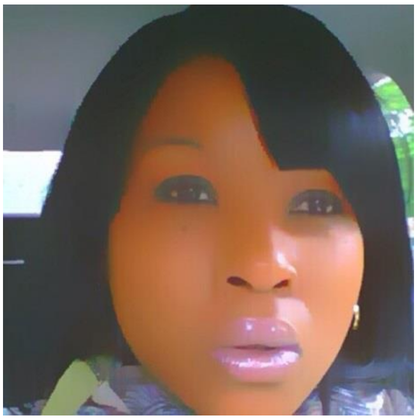


Figure 1-10: Instagram filtered photograph (Bossip, 2014)

These stories of image manipulation are repeated again and again, and our capability to perceive and be undeceived by such images remains confounded by the skills and easily accessible tools whereby the manipulations are created. So what do we do about it?

One hopes that we are able to be discriminating in viewing images to determine the additional meanings added or removed by the changes made to them. However, it is most probable that we are not as good at spotting manipulations and enhancements in images as well as we might like, as my research experiments discovered.

1.3 Digital image forensics vs. authentication

Assessing whether an image is manipulated can be considered from two different perspectives. The first perspective is digital image forensics, in which existing images are assessed retrospectively, using a variety of algorithms such as edge detection and within-picture pixel block comparisons to identify anomalies in the pixels of the images. The second, and the trajectory of this research, is proactive image authentication, a forward looking approach that seeks to establish and preserve the provenance and veracity (the authenticity) of images being produced now and in the future.

There is a marked difference in the amount of research that has been done in digital photograph forensics compared to authentication. Of the 30 plus key references regarding technological solutions to digital photo credibility listed in this paper, about 80% describe forensic research and 20% address authentication; while of course more could be listed from either category, the proportion in the literature remains significantly unbalanced. This appears to also be the case in research on manipulations in videos, another medium struggling with the enormous number of potential alterations that can be applied (Milani et al., 2012). A recent study stated that “passive approaches are the most important methods in the detection of digital video forgery” (Al-Sanjary p. 213) as opposed to active measures such as watermarking (Al-Sanjary & Sulong, 2015).

Forensic approaches can only conclude that an image has been manipulated, or that the image has probably not been manipulated. They do not focus on or incorporate the idea of positively asserting image authenticity and preserving that authenticity into the future. Further, digital forensic approaches do not seek to provide any insight into the meaning or significance of the image.

Current authentication research investigates ways in which the original or presentation version of an image can be preserved – these investigations are largely responding to commercial needs in protecting license revenue from images and tend to concentrate on various forms of steganography – insinuating visible and invisible coding into images. Early attempts at steganographic protection in digital image distribution have resulted in the ubiquitous ‘watermark’ to be found on images in many image licensing services. There have been some other approaches to authentication such as encrypted photo files and secure cameras, and these also are considered in chapter 2.

1.4 Summary and Scope

The field of image credibility is at present largely confused and ill defined, reactively responding to issues of image falsification with forensics, in preference to seeking authentication systems. Despite many excellent investigations into specific technological mechanisms for detecting photo manipulation, such mechanisms generally begin with 'sourceless' images whose provenance is already unknown and apply techniques that can only predict likelihood of possible manipulation. The field of proactive authentication is still relatively small, and few solutions have been attempted.

Currently the social recognition of the import of digital image credibility problems is still quite low, which counter-intuitively is actually beneficial. This is because the entrenched perception of the photograph as evidence of the real world has not yet been as devalued as might otherwise be the case in the face of the proliferation of manipulated images. This allows for authentication solutions that may be developed in the near term to be introduced while photographs are still widely regarded as accurate representations of actual people and places rather than the hodgepodge of truth and fiction they have actually become.

Past photographs are already entrenched in their own technologies and eras, but there is much we can do to preserve the future credibility of our photographic images. To address the current issues of misleading images, we need to turn to some of the same technologies by which they are created.

The initial intent of this research had been fairly straightforward, to investigate and build a proposed new file format, a mobile self-contained image format (.msci). However, upon the conclusion of Experiment A in which elements of this file format had been trialed together with investigations into human perception of manipulated and unmanipulated images, it became clear that that the ability and interest of viewers in interacting with and understand both manipulated and unmanipulated images was complex. It became necessary to better understand human perceptions of these manipulated and unmanipulated images as an important step in understanding the role they played in the usefulness of any new image format or authentication framework. Consequently, the subsequent experiment delved further into these human factors to better inform the resulting design of the image authentication framework.

As a result, this research addresses areas of image authentication that have as yet received little attention – investigating human perceptions of image credibility, and proactively presenting images within an environment of credibility that allows us to know what we are looking at and how it relates to the original image.

Overall, the research investigates image credibility from a human-centred computing (HCC) perspective. Using HCC approaches, I investigate if we as humans attend to the implications of image tampering, with or without prompting, and whether we can identify when we are viewing images that have been manipulated, and what those specific manipulations are. It explores how we use critical thinking about image veracity. When given tools to compare original and tampered images, is our ability to detect image interference improved? Are such tools enough? Is it possible to develop a human-computer framework that robustly supports image veracity?

The investigations and experiments undertaken in this research focus on discovering how well humans see manipulations in images using eye gaze tracking, and how they perceive them using questioning and verbal responses. In addition, a survey of attitudes towards manipulated images was undertaken throughout the research. Using neural network pattern matching, I investigate whether there are distinguishable patterns in eyegaze that differ between views of manipulated and unmanipulated images. Finally, I drafted a design specification for an image authentication framework based around a proposed new file format, and some elements of this file format were built and used in one of the experiments.

Importantly, this research takes a cross-disciplinary approach to investigate not just what technology is required to secure our images into the future, but how that technology is informed by our perceptions of and attitudes towards image manipulation and the roles such images play in human interaction and society.

This cross-disciplinarity can be envisaged across the range of research activities comprising this thesis as mapped in Table 1-1.

Table 1-1: Cross-disciplinary research

<table>
<tr><th>Human perceptions and attitudes</th><th>Information and communication technologies</th></tr>
<tr><td colspan="2">Snap survey of attitudes toward manipulated images</td></tr>
<tr><td>Attitudes towards manipulated images
Ability to perceive manipulations in images</td><td>User attitudes about need for technological solution</td></tr>
<tr><td colspan="2">Experiment A</td></tr>
<tr><td>Attitudes towards manipulated images
Ability to perceive manipulations in a range of manipulated and unmanipulated images
Our interpretations of manipulated images
Whether support tools assist us to see manipulations in images</td><td>Eye gaze tracking to identify visual cues and ability to see manipulations in images
User testing experimental image packaging modes.</td></tr>
<tr><td colspan="2">Experiment B</td></tr>
<tr><td>Ability to be trained in perceiving image manipulations
Separately comparing human perceptions of manipulated vs original images
Understanding meta-narratives in human perceptions of manipulated images</td><td>Eye gaze tracking in different subject cohorts exposed to differentiated images
Eye gaze tracking to identify visual cues and ability to see manipulations in images vs original versions
Grounded theory to tease out meta-narratives</td></tr>
<tr><td colspan="2">Neural Network Analysis</td></tr>
<tr><td>Understand if human vision uses distinguishable patterns when viewing manipulated and unmanipulated images</td><td>Understand if neural networks can predict manipulation state of images and/or human decisions about that state.</td></tr>
<tr><td colspan="2">Image authentication framework</td></tr>
<tr><td>Propose image authentication framework to improve human understanding of image credibility</td><td>Use outputs of experiments as input to develop image authentication framework</td></tr>
</table>

2 Literature Survey

The literature relevant to this research spans the following areas:

- nature of image manipulation,
- investigations into human perceptions of photo manipulation,
- eye gaze tracking and manipulated images,
- neural networks and manipulated images, and
- known efforts to create authentication solutions addressing the problem of preserving unadulterated photographic representations and the meanings they hold.

2.1 Nature of image manipulation

2.1.1 The rise and rise of photography and concomitant photo manipulation

In the 150 years commencing with the invention of conventional photography by Louis Jacques Mande Daguerre via announcement to the French Academy of Sciences in Paris 1839 (Bajac, 2002) and Henry Fox Talbot (Talbot, 1839) and long before digital photography was introduced, photographers have been creating important records of historical people, places and events, and staging images or creating seemingly real but actually false images crafted from disparate negatives in photographic darkrooms. Digital photography, introduced in the late 20th century, spawned 'digital darkrooms' to which photographers' photo manipulation activities easily transferred and extended.

2.1.1.1 The birth and early years of photography

Prior to the 19th century the chief method of recording images of people and places was through drawing and painting. This rapidly changed with the advent of photography in 1839, when Daguerre and Talbot introduced silver halide photography to the world.

Early photography was characterised by a wide range of photochemical techniques and it is beyond the need of this thesis to explicate them all. However, as an indication of the technical skill required to undertake early photograph production, it is useful to recount one of the more long-standing processes: wet-collodion photography.

To create merely the *negative* image using the wet-collodion technique popular in the mid-19th century, photographers had to make a syrupy collodion by dissolving gun-cotton (ordinary cotton soaked in nitric and sulfuric acid and then dried) in a bath of alcohol, ether and potassium iodide. Then, they had to follow the process described by George Baldwin in his book *Looking at Photographs*:

> *"In the wet-collodion process, collodion was poured from a beaker with one hand onto a perfectly cleaned glass plate, which was continuously and steadily tilted with the other hand, to quickly produce an even coating. … When the collodion had set but not dried (a matter of some seconds), the plate was sensitized by bathing it in a solution of silver nitrate, which combined with the potassium iodide in the collodion to produce light-sensitive silver iodide. The plate in its holder was then placed in a camera for exposure while still wet … After exposure, the plate was immediately developed in a solution of pyrogallic and acetic acids. … When enough detail became visible … the negative was removed from the developer, washed in water, fixed with a solution of sodium thiosulfate to remove excess undeveloped silver iodide, and thoroughly washed to remove the sodium thiosulfite, and dried. With an addition of a protective coat of varnish, the negative was ready to be used to make prints."*
> (G. Baldwin, 1991)

This was an extraordinary feat, required for each and every negative, and but a preliminary step in creating a printed photograph.

It is tempting to consider this complex process through the lens of modern digital photography and Pinterest and the voluminous image production we experience today, but photography had a rarefied role in ordinary lives of the nineteenth century.

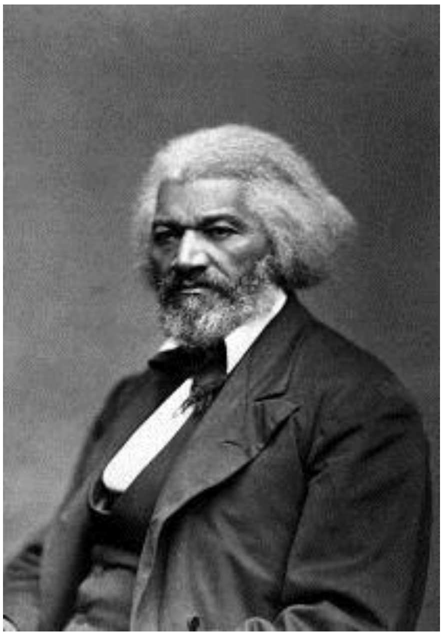


Figure 2-1: Frederick Douglass 1818-1895

Consider how many photographs any one person from the mid-1800s expected to own in their lifetime. Given the high cost and logistical issues (adequate dress, ability to access photography studios) many families might aspire to only one or two.

Even the most photographed person of the 19th century, abolitionist Frederick Douglass, who frequently sat for photographs to provide a counter-example to the extant stereotypes of African-Americans in the 1800s, can boast only 160 surviving photographs (Gregory, 2016).

Compare such a stellar record (by nineteenth century standards) to today's prominent individuals like Barack and Michelle Obama for whom over 400 images can be seen on the first page of Google Images alone.

If it was so difficult to create a photograph in the mid to late 1800s, then one might well understand making the argument as Henry Peach Robinson did that it was imperative that a photographer use all his tricks to make it a good one, including photo manipulation. Henry Peach Robinson is famous as an early practitioner of photo manipulation. His image *Fading Away* was made by cutting out and combining 5 negatives for 1 print in a montage. Although we know this from his description of it to the British Photographic Society in 1860 (Robinson, 1860), and from the same source, that the joins were placed in inconspicuous locations in the image, we do not know how the image was composed from the reported 5 negatives.

Figure 2-2: Fading Away by Henry Peach Robinson

Photo manipulation was key to Robinson's photographic practice. In his essay *The Legitimacy of Skies in Photographs* Robinson argued vigorously in favour of the practice swapping out a 'poor' sky for a 'good' sky in an image:

> *"It must be remembered that nature is not all alike equally beautiful, but it is the artist's part to represent it in the most beautiful manner possible; so that, instead of its being death to the artist to make pictures which shall be admired by all who see them, it is the very life and whole duty of an artist to keep down what is base in his work, to support its weak parts, and in those parts which are subject to constant changes of aspect, to select those particular moments for the representation of the subject when it shall be seen to its greatest possible advantage."* p. 61 (Robinson, 1869)

The same techniques used for the betterment of the aesthetics of an image were just as facilitative of image tampering. A well-known photograph of US President Abraham Lincoln is a fake. It was manufactured by an anonymous photographer who conflated Lincoln's face from a photograph by Matthew Brady with a standing portrait of earlier U.S. Vice-President John C. Calhoun (Aileen Jacobson, 2001).

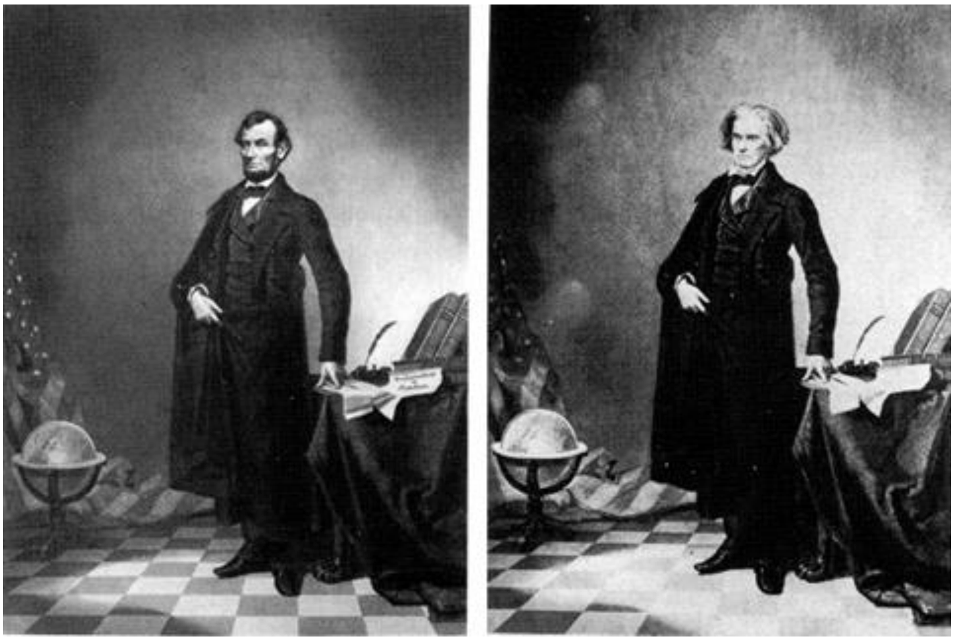


Figure 2-3: US President Abraham Lincoln – image on left is Lincoln's head superimposed onto John C. Calhoun's body (depicted on right)

Using a double exposure photographic technique lent itself well to ghost photography in the 19th century. In a time when photos were few and transport slow, it was common for families to have a post mortem photograph taken of their loved one, where the deceased was dressed up and arranged in as life-like a position as possible.

It was only a small step of imagination for photographers to create double and triple exposed negatives to join the 'ghost' of a deceased person together with a family member to create was became an entirely new genre of 'spirit' photography. This type of photo manipulation depended on the public's lack of knowledge of the chemical processes of photography – to many photography was mysterious and it was natural that a mysterious process would be able to capture the mysteries of spirits. Spirit photography lasted well into the next century until it was increasingly debunked in public forums.

Figure 2-4: Ghost of child visits. William H. Mumler *carte de visite* 1870-75 albumen silver print

The main types of photo manipulation of the early years of photography were combining negatives and multiple exposures, both of which were a form of splicing.

2.1.1.2 Coming of age

"You press the button, we do the rest."

Until the late 1800s, photography remained in the hands of the relative few, largely professionals with knowledge of the complex processes and chemicals required to produce a photograph. That all changed in 1888 when Kodak introduced the box camera and the 'snapshot' to the public. The camera came loaded with enough film for 100 photographs; customers had only to press the button on their camera until the film was exhausted, then send the camera to Kodak for refilling with film and eventual forwarding of the prints. This innovation transformed the photographic industry, removing the professional photographer 'middle man' and putting photography in the hands of the public (Jenkins, 1975).

With the dramatic expansion of the photographer base came an equally dramatic expansion of photographs as an expressive medium, not just a representative one. In addition to the army of people taking photographs of places that had never before seen a photographer, there were many who were experimenting with the medium.

Figure 2-5: Staged photographs - Cottingley fairies 1920 and the Loch Ness Monster 1920

Staged photographs made their appearance during the early 20th century. Two famous examples are the Cottingley Fairies and the Loch Ness Monster. In these cases, technically the images are not manipulated. It is the context and story behind the images that is contrived to create a false meaning. The Cottingley Fairies photograph was the brainchild of two young cousins Frances Griffith and Elsie Write, who photographed Frances with cutout sketches inspired by the *Princess Mary's Gift Book* depicting fairy figures, and presented them as 'real' fairies (Owen, 1994). The Loch Ness Monster came into being when two locals near Loch Ness convinced the town's doctor to take a photograph of a child's toy and to crop it to make its scale unidentifiable. They then spread the rumour of a serpent monster in the lake that is still alive today if modern photographic fakes are anything to go by (Telegraph UK, 2013).

But image manipulation was not just occurring as curiosities and art, it was also being used effectively to meet the agendas of people in power. A well-known example of this type of political history tampering through photographic alterations is the campaign of 'erasures' by Joseph Stalin (Figure 2-6), who had a habit of deleting political comrades

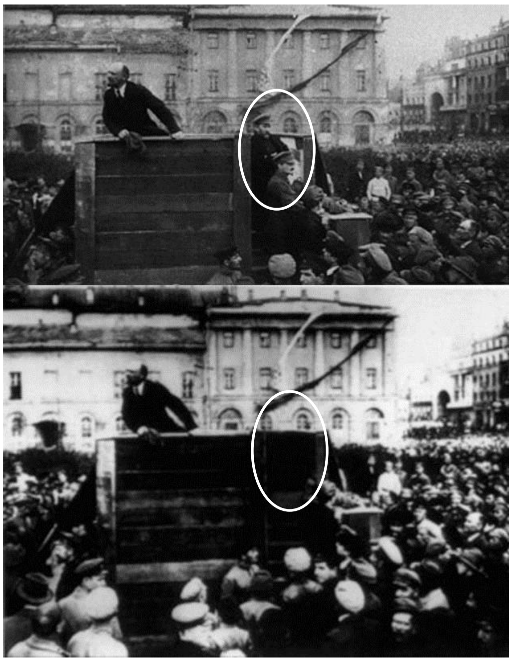


Figure 2-6: Stalin has Leon Trotsky and Lev Kamenev erased from the photo of his 1920 speech in Sverdlov Square, Moscow

from his photographs whenever he had deemed it necessary to delete them from physical existence by assassination.

Although photo manipulation still required professionals in the darkroom, often with expensive equipment to effect, there was money to be made in manipulated images. The advertising industry quickly adopted photography and photo manipulation as powerful sales tools.

Nowhere is the problem with image manipulation as prominent in the public mind as in the use by advertisers and print media of images of women slenderized and idealized with air-brushing and image tampering. Thin models in magazines have been implicated for decades in eating disorders of young girls, and poor body image of women generally as they compare themselves to images of models that have been altered out of all normal proportions and imperfections (Groesz, Levine, & Murnen, 2002).

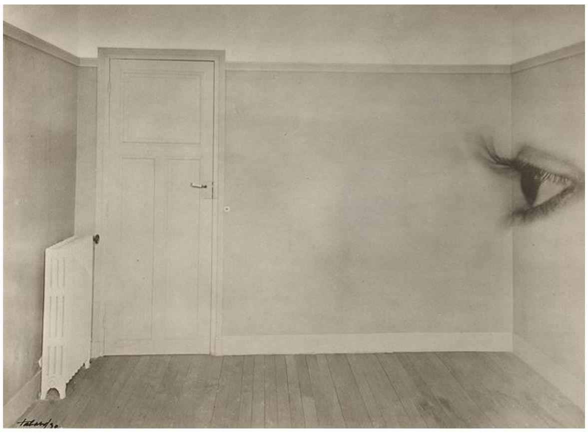


Figure 2-7: Room with Eye. Maurice Talbot, 1930

On the lighter side, images were (and are) manipulated for art production, one example of which is Maurice Talbot's Room with Eye (Figure 2-7).

For the next century, cameras would become steadily more functional, with the introduction of everything from the ubiquitous 'Brownie' (Olivier, 2007) to sophisticated SLR cameras to underwater disposable cameras.

There was however one major exception to this pattern of user-friendliness: photographs were still based on a light-sensitive chemical process that required film to acquire images, and either access to a film developing darkroom or (more likely) a supplier of photo processing services who developed the film and created prints from the film negatives. Fortunately there were plenty of companies providing these services; for example, in the USA during the 1980s, the film processing company Fotomat alone had over 4,000 drive-up kiosks (Skene, 2014), on concrete islands in parking lots. Photography had come of age.

Double exposing film negatives was still possible with modern cameras because for most of the 20th century camera film had to be advanced manually, and double exposures often occurred accidentally due to the photographer forgetting to advance the film to the next frame. Double exposing a negative on purpose to create photo effects was trivial. As noted earlier, air-brushing was rife in industry, and splicing negatives or double exposing prints to create montages were used in industry and art. An example of using spliced negatives in art is shown in Figure 2-7.

2.1.1.3 Photography in the Digital Age

Although some working prototypes of digital cameras existed prior to the 1980s, digital cameras were introduced in a commercial way with the Sony Mavica system in the 1980s, and became widespread in the late 1990s (Kawamura, 1998). Digital cameras offered two significant benefits over film photography. First, although the entry cost of a digital camera at that time could be quite high, once the camera was purchased there were no ongoing costs such as film and print development to constrain the photographer.

Perhaps more importantly, together with digital cameras came the unprecedented ability to instantly view the photographs taken, frequently while the scene was still available to be taken again to correct errors. Photographers no longer needed to seek a dark place to load a roll of film that would produce a maximum of 36 images whose quality would not be seen for days, if not weeks or even months. Instead, photographers inserted a media storage device and captured as many images as they wished, adjusting their photographic techniques as needed on the spot. Production of photographs boomed, with photographers producing hundreds of images a day if they so desired. Since that time, digital cameras have diversified into everything from high end digital SLR cameras, to cameras in everything from phones to iPads, to cameras embedded in USB hard drives, key chains, and Lego blocks.

As a result, the production of photographic images is at a level not seen at any time in the past. To illustrate the sheer volume of modern photo production, artist Erik Kessels created an art installation in which he printed out a copy of every photograph uploaded to the photo sharing site Flicker in a 24 hour period in 2011 (Newstex, 2011) and piled them in drifts in the Foam Art Gallery (Figure 2-8). In 2011 that figure was about 1 million. Users of Facebook alone upload more than 350 million photos every day (Facebook, 2013). According to the business data company Domo, Pinterest receives about 200,000 images per hour, and Instagram users post over 200,000 images per minute (Domo, 2016).

Figure 2-8: Erik Kessels, 24 hours of Flickr, 2011

Digital photography and the long arm of the Internet increased the problem of photo manipulation. Photographs, now ‘born digital,’ began life as a digital file, and software to edit that image file was then, and continues to be, readily available. Most notably, Adobe Photoshop, developed in the late 1980s to early 1990s, has become synonymous with the act of image tampering: we say a photo has been ‘photoshopped’ to mean that it has been altered. Microsoft has for years been seeking probabilistic solutions for automatically using the background of a photograph for filling in the ‘holes’ left when an element of the photograph has been deleted, such that the unwanted element simply ‘vanishes’ (Financial Times, 2002).

Image manipulation software has become inculcated into photographer’s post-processing of photographs, offering easy access to an extensive palette of image tampering tools. Such manipulated images are now common; although often manipulated for fun or art, just as in the past some photos arere manipulated for political ends.

The paper “Digital image forensics: a booklet for beginners” describes a recent example of this:

> *“In July 2010 Malaysian politician Jeffrey Wong Su En claimed to have been knighted by the Queen Elizabeth II, as recognition for his contribution to the international aid organization Médecins Sans Frontières. A picture of him being awarded by the Queen of England accompanied his statement, diffused in local media (a). When questioned about the award though, the British High Commission in Kuala Lumpur made clear that the name of Mr. Wong was not included in the official knighthood recipients lists, and that the picture was inconsistent with the usual protocol adopted for knighthood ceremonies. The image was finally shown to be a splicing between an original ceremony photo (b) and Mr. Wong’s face, built to increase his popularity.”*(Redi, Taktak, & Dugelay, 2011)

Figure 2-9: Jeffrey Wong Su En being 'knighted' 2010, spliced image on left

Most notable of the many errors in protocol that present in this image are the fact that a person receiving a knighthood is ‘dubbed’ by the Queen as he kneels on an investiture stool; he does not get a decoration on his jacket (Royal Household at Buckingham Palace, 2016). This means that the original photograph was not of a knighthood ceremony, and Wong Su En’s image is both manipulated *and* misrepresentative.

In response to the ease with which digital images can be altered, a range of digital image forgery detection techniques have been developed in recent years (described further in section 2.6.1). Collectively these techniques are known as digital image forensics, a field that analyses images to determine image veracity through identifying image manipulation artefacts. Interest in digital image forensics techniques is increasing due to the potential of image manipulation to impact on medicine, justice, news reporting and the legal and accounting professions.

2.1.2 Types of photograph manipulation

While many of the principles ascribable to photo manipulation can be more broadly applied to images of every type (vector graphics, diagrams, art images), as noted in the preamble, this research focuses on photographs captured by cameras as physically-based reflections of the physical world (photographic images), and it is this reflective relationship that is arguably both most valuable to society and simultaneously most vulnerable to loss in image manipulation.

In the previous chapter we saw several high-profile examples of image manipulation. But what exactly is image tampering in the digital era? What are the specific techniques whereby such tampering occurs? The terms used to describe image manipulation may vary but the actual techniques used are well known, all of which make use of the fact that images can only be displayed for viewing in the form of a matrix of pixels (picture elements). Image tampering techniques therefore seek to change pixels in images.

There are many ways in which pixels in photographs are altered from the images produced by a camera. Some of these could be seen generally as innocuous: any portrait photo taken by a digital SLR camera must be rotated 90 degrees in order for it to appear upright unless it is simply printed out onto photo paper from the original photo file. Straightening a ‘crooked’ image also involves rotating the assemblage of pixels even if it is only 1 or 2 degrees, but invariably involves the risk of lost information by cropping to neaten the edges. Removing the unwanted photographic effect ‘red eye’ in which flash photography highlights the red colour of the blood vessels at the back of the eye, is so common that it is a standard function in image processing software and even in cameras.

There are also many ways in which settings in a camera can alter a photographic image before it is delivered to the memory storage device; these photographs may be enhanced before they are properly recorded. Although this is largely beyond the reach of my research, it is worth noting that such alterations should at least form part of the metadata, and that such alterations should be reversible, or contained within a separate image file.

There are a range of global changes - such as colour correction, resizing, and adjusting brightness - that may make little or no difference in the meaning in the image. However, global changes still have the potential to impact significantly upon the pixel data. For example, an image that has been multiply resized and saved can degrade to an unacceptable quality (Fig 2-10). This image degradation can be used to hide image manipulation.

Figure 2-10: Photo of infant in arms – resized x1 (left) and resized x10 (right)

It is the more invasive manipulations that are the most concerning, tampering that allows for alterations in images so significant that, depending on the extent of alteration, the meanings in the images can be transformed. These are splicing, copy/move and retouching.

2.1.2.1 Splicing

Image splicing techniques are used to compose one image from multiple images. The earlier 'doctored' images of Jeffrey Wong Su En being knighted by Queen Elizabeth II and the Abraham Lincoln head swap are examples of this type of manipulation. In splicing, multiple images are involved. One image is usually designated the main

Main image, Floriade 2006

Secondary image, Floriade 2006

Figure 2-11: Spliced image shows all three people together in same photo

image, and elements of other photographs are outlined and copied, then pasted into the original image. In Figure 2-11, one photograph at the Floriade flower show in Canberra, Australia has been augmented with another person from a second photograph.

Detecting splicing is difficult. The current digital forensics focus in these cases is investigating where the changes show up best: the places where the outermost pixels of the spliced region come into contact with the original pixels in the image. The presence of sharp edges (or changes in linear pixel values) in this area can help reveal the fact that these image fragments came from two different photos. These detectable changes are not proof against counter-measures. For example, determined forgers can apply one or more of several types of filters that ‘smooth out’ the abrupt changes in pixels (Kirchner & Böhme, 2007). This means that forensics detection measures are not reliable now, or into the future.

2.1.2.2 Copy/move or cloning

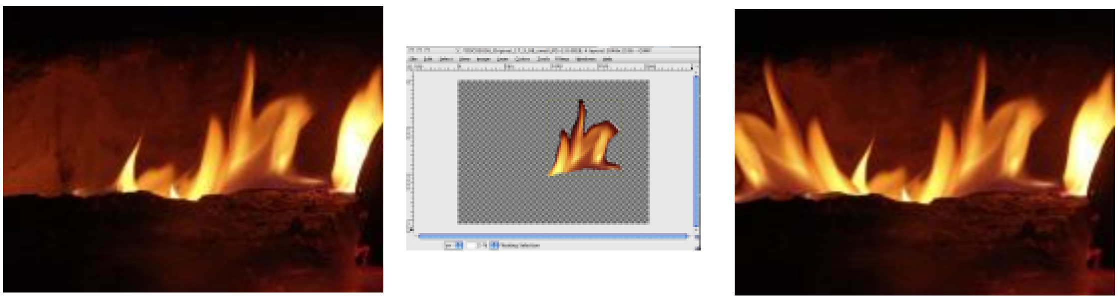


Figure 2-12 Original campfire photo (left) Flame copied (middle) Copied flame flipped, moved and pasted in (right).

Copy/move (or cloning) forgery is a popular form of tampering, in which a target region is copied from a particular location in an image and thereafter pasted at one or more locations within the same image. Because the target region is copied from and pasted into the same image, the colour and brightness values of the target region meld easily with other areas of the image and the resulting altered image can appear completely natural to the naked eye (Fig 2-12).

2.1.2.3 Image retouching

Image retouching is about making slight changes in a photo for various aesthetic and commercial purposes. The retouching is mostly used to enhance or reduce the image features. This used to be called air-brushing because prior to the digital age retouching of photographs was done with paint using handheld airbrush paint guns. Today this can be done easily with Photoshop, as can be seen in Figure 2-13 (Dolly, 2009).

Figure 2-13: Popular TV personality Jessica Mauboy; retouched image on left has airbrushed out the two blemishes on cheek and forehead visible in image on right. Her neck has been lengthened, and her jawline has been reduced, making her face slimmer and smile unrealistically large.

So far, all image modifications belong to one or more of four basic manipulation techniques: splicing, copy/move, retouching, and global changes. Table 2-1 indicates how common image manipulations are usually effected through these four main types of image manipulation.

Table 2-1 Image manipulation effects by category

Action	Splicing	Copy/move	Retouching	Global changes
Collage / photo montage	✓	✓	✓	✓
Adding/deleting elements	✓	✓		
Retouching / airbrushing			✓	
Colour adjusting			✓	✓
Brightness adjusting			✓	✓
Red-eye reduction			✓	
Rotating				✓
Cropping				✓
Resizing				✓

2.1.3 Factors outside of the image affecting credibility

The representative nature (or otherwise) of images is not solely determined by their contents. From its inception, every photograph is guided by the will and ideas of the photographer. What ultimately is captured within the frame may be serendipitous or carefully crafted, indeed, staged. Further, images are not normally viewed on their own. They are surrounded by text, colours, graphics, even advertisements (in the case of online environments like social media). Digital images are normally accompanied by metadata even if many people are not aware of its existence or how to access it. These elements, staging, context and metadata, influence the credibility of the image we see and our ability to understand it.

2.1.3.1 Staging

Credibility but not authenticity is affected in staging of photographs. There have been numerous examples of this type of image forgery in the 19th and 20th centuries. In addition to the Cottingley Fairies and Loch Ness Monster photographs mentioned in 2.1.1.1 there have been many photographs staged to mislead the viewing public. There are several reasons to modify an image by staging, which range from attempts to capture meaning not otherwise readily gleaned from a landscape or event, through to less beneficent motivations. Conveying the sense of a scene or event in a single image is one of the more frequently quoted reasons given by photographers who do this. Examples of this intent can be seen in the iconic image of the Spanish Civil war by Robert Capa, The Falling Soldier, and the Brady/Gardner staged American Civil War photos.

Figure 2-14: The Falling Soldier, 1936, Robert Capa

The Falling Soldier photograph, it is now known to be staged in contradiction to Capa's original title of the photograph, "*Loyalist militiaman at the moment of Death, Cerro Muriano, September 5, 1936*." It has been proven from geographical features of the image that the photograph was taken in Espejo, Spain, 48 miles away from the battle at Cerro Muriano. Since then, two previously unknown photographs of staged falling soldiers have been found amongst Capa's negatives (Rare Historical Photos, 2016).

Matthew Brady and his partner Alexander Gardner were well known to stage photographs during the American Civil War. Often bodies that littered battlefields were too far apart or not positioned properly to achieve the effect the photographers desired. Brady and Gardner dragged bodies together, positioned bodies, and arranged artifacts like rifles and clothing to maximum effect. Figure 2-15 shows an example where the bodies were not just grouped, but actually moved *en masse* to create a staged image of war dead in front of a church.

Figure 2-15: Dead soldiers moved to foreground of Dunker church, Battle of Antietam

Ironically, a photograph of a staged scene is still authentic, albeit an authentic photograph of a non-credible scene. Even though the scene was staged, it is in fact the real scene recorded by the sensor of the camera, staging and all.

2.1.3.2 Context

Photographs can be unmanipulated and still be untrue. This is especially true of staged images as we saw above. However, in addition, photographs often appear within an envelope of textual or verbal information, which can misrepresent what is portrayed in the image.

Examples of this type of misconduct arise in every field. In 2006 the Editor-in-Chief of a highly cited biology journal, the Federation of American Societies for Experimental Biology (FASFB) journal, noted an example of scientific fraud where simply changing the context of the image falsified its meaning: a Norwegian oncologist who used the same micrograph at different magnifications as representative to two different patients and disease stages. (Weissmann, 2006).

Figure 2-16: 'Lost Amazonian tribe'

In June 2008, photographer Gleisen Miranda took photographs from overhead of a tribe in the Amazon. They were presented to and released by the media (including The Canberra Times, Chicago Sun-Time and the BBC) as photographs of a lost tribe, when in fact the tribe was already known, and the photographs were being used as an argument against logging in the area. The photographs were not faked, but their context was.

In other contexts, images can get lost in inadvertent campaigns of misinformation. With haphazard search algorithms, the go-to information source for much of the public, Google, can and does present images in potentially misleading contexts, and images can be presented incorrectly by often well-meaning but ultimately untrustworthy online sources. A detailed example of this type of contextual misinformation is provided at Appendix B, The Case of the Venezuelan Poodle Moth.

Figure 2-17: Venezuelan Poodle Moth

2.1.3.3 Metadata

Metadata is an information layer that co-exists with a data resource to describe or explain the nature of the data source. It is often referred to as 'information about information' but a more useful definition was written by the US National Information Standards Organization (NISO) in 2004 that states that metadata is "structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use or manage an information resource"(NISO, 2004). It is usually based on the MARC standard (MAchine Readable Cataloging tags) that has been in use in one form or another since the middle of last century (Jackson, Lubas, & Schneider, 2013)

The metadata captured by many modern digital cameras is rich and useful, providing details about camera settings, pixel dimensions, image compression etc. (Figure 2-18) and in some cases can be augmented by user input data for even more clarification about the meaning of the image.

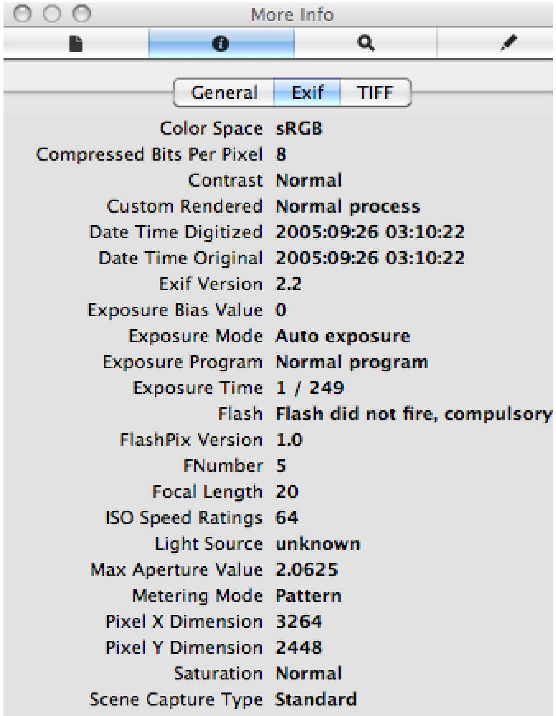


Figure 2-18: Sample camera metadata (Sony DSC828)

Unfortunately, in many cases, including the case of digital images, metadata is not immediately visible, and must be navigated to for display. Adding additional metadata information such as names of people in the images, or identifying the event is a process that many people do not even know exists, let alone how to use.

Metadata is a powerful adjunct to the content within digital images to ensure their continued intelligibility. As NISO remarks, "metadata is key to ensuring that resources will survive and continue to be accessible into the future" (NISO, 2012, p.1). Finding ways to make metadata more user-friendly can help increase the likelihood that knowledge about future images will be preserved.

2.2 Human perceptions and perspectives of photo manipulation

There has been significant research on how we see, and how we view images. We know that human vision occurs as an alternating sequence of fixations and saccades. In fixations we fix our gaze upon a small area of the visual landscape. Saccades, discovered by Louis Émile Javal in the late 19th century, are the motions by which we move our fixations from place to place. While our eyes are in the saccade state we experience brief (milliseconds) vision gaps (N. J. Wade, 2003). The mechanics of this vision process are heavily augmented by our cognitive processes through both the control and interpretation of the visual fixations.

Examinations of saliency in features of an image have shown that we focus on salient features of an image to discern its meaning (Itti & Koch, 2000), (Treisman & Gelade, 1980), and investigations of eye gaze scan paths demonstrate that some characteristics of images, especially luminance (Harding & Bloj, 2010) attract our attention.

There is also research on how we look at images. We look at images differently if we are looking at the characteristics of the image, the people in the image, or if we have specific reasons for looking at the image (Buswell, 1935).

However, there has been very little research to date on how we do or do not perceive manipulations in images, despite it being of interest to understand how humans interact with the myriad digital images of unknown veracity presented to us each day, and it is also of interest to understand to what extent we can discern ‘faked’ photos, and how we interpret the meanings presented in them.

Of the few studies in this area, it seems evident that humans are poor at noticing image manipulations. Concerned with the ability of observers to make perceptual judgements about scene geometry in making a distinction between real and fake images, Farid and Bravo conducted a study examining this judgement and found that “except for the most degenerate cases, performance was near chance, even though the information required to make these judgments was readily available in the scenes” (Farid & Bravo, 2010).

We find it difficult to spot image tampering even when the images change as we view them. In one study, it was shown that even quite significant changes in scene (exchanged heads on two cowboys on a bench or a swimsuit colour change or a building suddenly becoming 25% larger) were not detected by experiment participants more than half the time, especially if the change occurred during saccades rather than fixations (Grimes, 1996).

This lack of research specifically into how we can or cannot perceive manipulations is surprising, because our ability to discriminate between the truth and falsehoods presented to us in image form is important. Images are consumed by humans in a range of contexts, sometimes with differing motives from those presenting them to us. They may be casually observed, or intently studied. We may understand them to be manipulated or not, and all of these perspectives factor into our received meaning from the images.

On a moment-to-moment basis, our skewed perceptions as we absorb these manipulated images may not be consequential. But our views, choices and actions can at times be based on the information we receive in visual form, especially photographs. For example, it has been shown that US presidential countenances can be enhanced using judicious image manipulation (Keating, Randall, & Kendrick, 1999). It is well-known that women and girls' perceptions of body image have been significantly impacted by comparing themselves to the idealized women who have been trimmed, airbrushed, and polished in the press. It is this area that has attracted the most attention in respect of image manipulation. Concerns with the presentation of girls and women in the popular press have been expressed for decades. Perhaps as a foreseeable consequence, airing these concerns has created a more knowledgeable consumer base, who are rejecting these offerings. (Reaves, Hitchon, Park, & Yun, 2004).

The reason people manipulate images is precisely this: to alter our perspectives on the information being displayed in the image from what we might be expected to perceive from the original, to what it is desired that we perceive from the altered version. The motivations for image tampering and thus changing our perceptions of an image may be beneficent, benign or maleficent.

2.3 Peak photographic organisations' responses

The problems inherent in heavily edited, even fictitious images have been attracting public attention and discourse (for example the controversy of *Gaza Burial*, the World Press Photo of the Year 2013).[1] With the growing recognition of these issues, beneficial changes are beginning to appear. One such change is that the window of opportunity for photograph manipulators to submit altered images to photo competitions as 'real' photos is closing. It seems that from 2015, photography competitions expect photo entries to contain less fiction and more reality.

[1] Although the photograph was not really faked, it was tremendously 'photoshopped,' unfortunately overshadowing the real human tragedy captured in this important photograph (Steadman, 2013).

For example, in 2015, as part of their annual Canon Light Awards, Canon announced that entries must be actual photographs, not manipulated images. Specifically, they respond to the question "Will my entry be valid if I use photo editing software (like Photoshop)?" with the following statement:

> *"Yes, basic editing such as cropping and colour adjustment is permitted. However, entries must be true photographs and not composites or digital manipulations. Keep in mind to confirm you are a winner, you may need to send through the original file …" (Canon, 2016)*

National Geographic is very specific about what types of photo editing is allowed. Before detailing what is allowed using techniques like burning and dodging, compositing, captioning and overall ethics, they expound their philosophy on manipulating photographs:

> *"Our biggest ask is that the photos stay true to your personal vision and to what you saw. Please avoid heavy-handed processing. We want to see the world through your eyes, not through the excessive use of editing tools. If the photograph is manipulated, please describe your process in the caption." (National Geographic, 2016)*

There are many more examples. The International Loupe Awards rules state of their photojournalism competition that "Winning images in this category deemed to be composited images will be stripped of their category placing, prize money and or prizes" (International Loupe Awards, 2016). The International Pano Awards is less convincing, but still notes that manipulation may lessen the photographer's chances: "Images may be from single capture or stitching software, film or digital capture, but must be 100% photographic in origin. Manipulation is allowed but excessive manipulation may be scored down by judges" (International Pano Awards, 2016). The Smithsonian says of their photographic competition that "we do not accept digitally or otherwise enhanced or altered photos, except for those entered in the Altered Images category … If the judges determine that a photographer has altered his or her photo, they reserve the right to move the photo to Altered Images or to disqualify it" (Smithsonian, 2016).

World Press, the premier international photography competition for professional photographers, introduced rules in the World Press Photo Competition after the genuineness of the 2013 winning photograph, *Gaza Burial*, was questioned.

> *"Participants are now required to provide file(s) as recorded by the camera for all images that proceed to the final stages of the contest. These file(s) will be requested and studied confidentially during the judging period (1-11 February 2015). A failure to provide these files before 11 February 2015 will lead to the elimination of the entry."* (World Press, 2016)

Not only have World Press built anti-manipulation rules into their prestigious photography awards, World Press Photographs of the Year, in 2016 they are enforcing them. In this year, World Press Photo stripped photographer Giovanni Troilo of his Contemporary Issue World Press Award 2015 over Troilo's admission that one of his photographs on the theme of urban/human degeneration in Charleroi, Belgium was actually taken in Molenbeek, Brussels. According to Business Insider, "World Press Photo (WPP) said in a statement that the [award] had been withdrawn over "misleading information"."(Lefour, 2015)

World Press Photo withdrew Troilo's award. Giulio Di Sturco and Tomas van Houtryve, second and third place winners of the category were raised to first and second place. There was no-one elevated to the now empty third place position. Troilo's choice to mislead was not victimless. Consider a) the weight of false negative meaning Troilo added to his expose of the 'dark underbelly' of Charleroi, and the impact that would have on the general *esprit d'corps* of the people of the city of Charleroi, and b) think about the person who should have come in third in the World Press Photo competition, who did *not* win that prestigious and career-boosting award. Troilo shrugged off the debacle, stating that "people have been writing to me from Australia, from Switzerland, about exhibitions, because of the publicity" (Smyth, 2015).

Regardless, photographic competitions seem to be giving weight to the connection between the aesthetics and meaning of photographs and the real world they purport to interpret on our behalf. Yet despite this, many organisations still side-step the issue; in a recent '2012 Citizenship Report,' (Klein, 2012) Getty Images did not even mention digital photograph credibility despite extensively discussing other broad photographic issues such as intellectual property rights and copyright, press freedom and the protection of journalists and easy legal access to digital content.

2.4 Eye gaze tracking and manipulated images

Eye gaze is an important window on our perceptions of image manipulation. Whenever our eyes are open we are using our gaze in a controlled fashion to help us navigate our worlds, both literally and figuratively. As early as the middle of last century we had understood that the movements of our eyes were driven by our minds to serve our cognitive and perceptual goals (Yarbus, 1967).

We use gaze control to seek out visual information relevant to the task at hand. We can therefore use eye gaze information as a window into attention allocation and as a behavioural index of visual and mental processes (Henderson, 2003).

It is therefore surprising that there is not more research available on using eye gaze as a window into our cognitive processes around perceiving falsity in images. What little there is (noted in 2.2 Human perceptions and perspectives of photo manipulation, above) seems to indicate that we are not highly successful at discerning manipulations in images (Farid & Bravo, 2010). As previously noted, even heads being swapped between two figures did not occasion great remark on the part of experiment participants, especially if the change occurs between visual fixations (Grimes, 1996).

2.5 Artificial neural networks and manipulated images

As yet, artificial neural networks do not appear to have been applied to interrogating human eye gaze responses to photographs and manipulated images. I could not find any papers on this topic.

In respect of using neural networks with images generally there has been some research. Images can be deconstructed into matrices, and neural nets lend themselves well to investigating matrices. As early as 2000 neural nets have been used in understanding colour and line features in images (Inoue, Mitsukura, Fukumi, & Akamatsu, 2000; Jerebko, Barabanov, Luciv, & Allinson, 2000), Not surprisingly, using neural networks with images has increased since that time in both breadth and scope; a recent paper described watermark authentication of telemedicine videos using a neural network encryption decryption process (Agilandeeswari & Ganesan, 2016).

Neural networks may be able to discern the biases of people viewing images at a less than conscious level. This was suggested by Tamás Gedeon in his paper based on an experiment in which viewer preferences of Mondrian style art images were approximated using back-propagation neural networks (Gedeon, 2008). Specifically, in recording that the neural network analysis of the main subject's aesthetics showed a surprising underlying relationship with certain arrangements and proportions of yellow, Gedeon noted "This suggests that our technique is able to identify relationships which are not consciously available to users."

Gedeon's research was part of the impetus to apply neural network processing to the eye gaze of experiment participants in this research. In addition to using neural networks to identify if people look at manipulated and unmanipulated images differently, it is also of interest to know if neural networks can discern signals in people's eye gaze that indicate relationships between elements that may not be consciously recognised.

2.6 Known efforts to create solutions

In recent times several techniques for detecting changes in digital photographs have been developed. In large part these exploit the relationships of pixels in a photograph to one another. Changes in light intensity, abrupt local shifts in pixel colour values, and discontinuities of pixels enable forensics technicians to identify image splicing, in which elements of different photographs are combined in a single image, and tampering, in which a photograph is photo-processed to change local areas of the photos. There have been many approaches to detecting manipulated images as described below. Far less emphasis has been placed on proactively asserting the authenticity of an image.

2.6.1 Digital image forensics vs authentication strategies

As discussed earlier (1.3), forensic research investigates ways in which existing photographs can be examined for manipulation. This is the most familiar form of technological solution to image manipulation, largely because it is in practical use and has been for many years in intelligence agencies. Even the most unsophisticated viewer of photographs participates in this type of detection, by applying common sense to assessing the truth of visibly manipulated images or noticing that various shadows of elements in the photograph 'don't add up.'

Figure 2-19: Frank Hurley, Shackleton expedition 1915

For example, the famous Australian photographer Frank Hurley was well known for creating manipulated images in the Antarctic at the turn of the last century, and although no computer technology eyes were available to cast doubts upon them, these manipulations were easily spotted by interested lay people. In Figure 2-19 it is possible to see that the penguins in the foreground of the image do not cast a forward shadow while the sled dogs in the mid-ground do – a physical impossibility even in the Antarctic. Hurley later admitted as much, but said that it was impossible to capture the scope of the camp because everything was too spread out to capture in a single scene (McGregor, 2004). Interestingly however, human vision does not tend to focus on shadows (Jacobson & Werner, 2004; Porter, Tales, & Leonards, 2010) and therefore we do not immediately see shadow clues to an image's inauthenticity, so not everyone is able to distinguish this problem in an image.

Because many image manipulations are difficult or very difficult to detect, a range of sophisticated techniques have been planned and/or implemented in recent years. Edge detection techniques have been proposed by (Ng, Chang, Lin, & Sun, 2006), who stated that abrupt changes of pixel values predicted image splicing, and (Dong, Wang, Tan, & Shi, 2009), who analysed the discontinuity of image pixels to detect splicing. This is one of the most prominent forensics techniques and has been subject of extensive research (Sutthiwan, Shi, Su, & Ng, 2010). There are also methods for detecting softer changes in pixel correlation such as multi-size block discrete cosine transform and moment of characteristic function, which average the light values of varying blocks of pixels and compare them with adjacent and overlapping layers (Wang, Dong, & Tan, 2014). Farid added a focus on the specularities, (light reflections) of the eye to compare details of the reflected environment to this list (Lang, 2007).

Other types of photo characteristics that have been investigated for usefulness in forensics include chromatic data (Wei, Dong, & Tan, 2009) , geometric invariants (Hsu & Chang, 2010), and mathematical analysis, such as statistics ((Ng et al., 2006)) and even viewing image data in spreadsheets (Waldcock, 2010). It is interesting that some of these forensics techniques have been used in reverse. For example, Hsu and Chang have identified that each camera image sensor has its own inconsistencies that act like a fingerprint, which means that individual images can be linked to the camera that took them. This is of particular usefulness in identifying child sex abusers when offensive images can be linked to a particular camera, which can in turn be linked to a particular person. Another approach is shadow analysis in which suspect image elements can be confirmed as inconsistent with the host image by geometric identification of light sources (Kee, O'Brien, & Farid, 2013) or light strength (Ke, Qin, Min, & Zhang, 2014).

There are many problems with forensic tools, chief amongst which is fact that they cannot verify that an image is original, but only identify positively that a detectable manipulation has taken place. Some of the forensic analyses are known to be effectively nullified using forensic counter-measures, such as smoothing filters to foil edge detection in spliced images (Kirchner & Böhme, 2007). Another challenge these tools face is their ability to cope with the increasingly large file sizes of photographs (Richard & Roussev, 2006), though this is a challenge that any authentication tool would also face.

To ensure that photographers and online image companies are compensated for the use of their images, many image archives (Getty Images, Shutterstock, iStockphoto and hundreds of other archives) apply an image distorting or obscuring watermark to images displayed on their site, which are only removed at the time of purchase. These techniques have arisen from research by many (Radharani & Valarmathi, 2010) who suggested using photographic feature amplification to achieve an authenticating watermark.

2.6.2 Finding authentication solutions

As is evident from the described techniques above, to date the main focus of image authentication has been retroactive; a *post-priori* process in which various techniques are used to ascertain whether an existing image has been modified.

As mentioned above, although many of these techniques can arrive at a definite determination that a photograph has been manipulated, they are not able to arrive at a determination that a photograph has not been manipulated. The best that can be achieved is a statement that the image is unlikely to have been manipulated. This problem has been recognised in industry, with at least three camera manufacturers attempting solutions although all three ultimately failed. So, despite these attempts, no authentication process for digital photographs has yet been successfully implemented.

2.6.2.1 Nikon's Failed Authentication Model

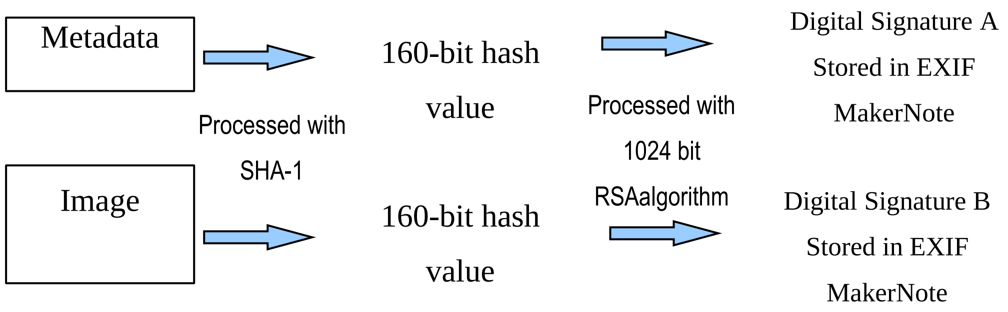


Figure 2-20: Nikon authentication model

In 2009 Nikon attempted a verification system based on hash signatures. Although Nikon is not vocal on the matter, the hacker organisation that broke the code, ElcomSoft, gave details of the system (Leyden, 2011) . Their research showed that image metadata and image data were processed independently with a SHA-1 hash function. There were two 160-bit hash values produced, which were later encrypted with a secret (private) key by using an asymmetric RSA-1024 algorithm to create a digital signature. Two 1024-bit (128-byte) signatures were then stored in EXIF Color balance MakerNote tag 0×0097.

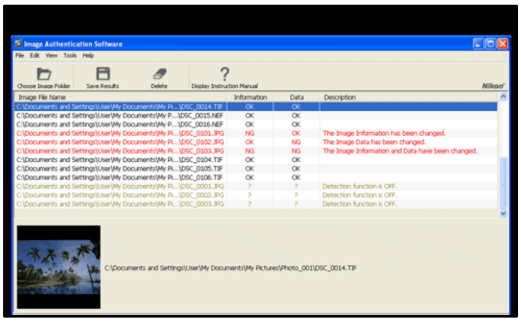


Figure 2-21: Screen capture of Nikon's authentication software output

There appears to have been three weaknesses in the system: 1) using SHA-1 to produce the hash value instead of SHA-2, 2) using an asymmetric security algorithm, which is considerably less secure than symmetric security algorithms, and 3) storing the digital signature in EXIF data where it can be rewritten by third parties interested in obviating the security feature.

The digital signatures worked with Nikon's authentication software (now unavailable) of which a screen capture appears at Figure 2-22 (screen capture was taken by Elcomsoft).

2.6.2.2 Canon's Failed Authentication System

Canon released an authentication system that was cracked in much the same way as Nikon's by Elcomsoft. The system, Canon OSK-E3 Original Data Security Kit was incorporated into the Canon 1D Mark III Digital SLR Camera.

Figure 2-22: Canon OSK-E3 Original Data Security Kit

It appears to have required a hardware device (USB connection) through which the images stored on the memory card of the camera were processed with Canon's encrypted security. Associated Press was using this system to authenticate images before it was hacked.

Canon described their system as

> *"The Canon OSK-E3 Original Data Security Kit for the Canon 1D Mark III Digital SLR Camera is a very secure way to make sure there is no unauthorized use of your files - either by previewing or pilfered on-line (wireless).*
>
> *This system, which requires the installation of included software for your computer (Windows), also comes with a USB card reader/writer that can verify the soundness of your data.*
>
> *In addition, the system has been designed to alert you to which parts of your data have been changed - such as pixels, etc. Imagine how safe you'll feel being able to encrypt your data which can only be decoded via an OSK-E3 system.*
>
> *For those shooting news, art, commercial work or anything that requires even a hint of protection, this system is well worth the cost." (Business Wire, 2007)*

It is difficult to identify the weakness in the system, but it seems the 'cryptoprocessor' allowed external access and reveals the secure key. One critic (Kirk, 2010) commented that Canon should implement HMAC calculation in their cryptoprocessor which does not expose the secret key. However SHA-1 used by Nikon is a HMAC and that did not stop Elcomsoft.

2.6.2.3 Epson's and Kodak's discontinued watermarking systems

Epson and Kodak have produced cameras with security features such as the Epson PhotoPC 3000Z and the Kodak DC-290. Both cameras added irremovable features to the pictures which distorted the original image, making them unacceptable for some applications such as forensic evidence in court. According to secure digital camera researchers Blythe and Fridrich, "[n]either camera can provide an undisputable proof of the image origin or its author" (Blythe & Fridrich)

Epson’s technique is the better one however, in that it was based on digital signature and private/public keys. Unlike Kodak's attempt, it can detect a single pixel change.

2.6.2.4 Theoretical Ideas of Secure Digital Camera (SDC) becoming reality

A secure digital camera (SDC) was proposed by (Mohanty, 2009), in 2003 and published in January 2004. This was not the only time this was proposed. Blythe and Fridrich also have worked on the idea of an SDC in 2004 for a digital camera that would use lossless watermarking to embed a biometric identifier together with a cryptographic hash (Blythe & Fridrich).

InformaCam – An initiative of Witness, currently in beta, developed in 2013 with initial funding of $320,000 from the Knight News Foundation, InformaCam allows users “to take images and videos, embed them with geotemporal and other metadata, sign them with a digital signature unique to the device’s camera sensor, encrypt and then send those files to someone they trust who maintains a secure server” (Velden, 2015). This has now evolved into a mobile phone image verification system for Android called CameraV. CameraV is a downloadable application that collects metadata about photographs and videos taken with the mobile and adding a unique hash key based on the camera’s specific pattern of sensor noise. It is a step in the right direction, but has limited functionality and may be subject to the same weakness that sank the Canon and Nikon attempts at an authentication framework.

2.7 Research goals

In light of the landscape of image manipulation and the current state of image authentication outlined above, this research has the following goals.

- Understand the attitudes toward and significance of the problem of image manipulation for lay people.
- Understand if people can perceive manipulations in images.
- Determine how people interpret manipulated images.
- Determine if people can understand image manipulations better when images are presented in a trial file format that presents manipulated images together with their original version.
- Learn if people can be trained to better discern manipulated images.
- Learn if computers can be trained to discriminate between eye gaze data for subjects viewing manipulated vs unmanipulated images
- Determine what people need in an end-user style technological solution to allow them to understand the images with which they are presented and their credibility.
- Design an image authentication framework that addresses known issues with manipulated issues (described in Chapter 2) and discovered in experimental investigations during this research.

3 Experiments

There has been significant research on how we view images. Examinations of saliency in features of an image have shown that we focus on salient features of an image to discern its meaning (Itti & Koch, 2000), (Treisman & Gelade, 1980), and investigations of eye gaze scan paths demonstrate that some characteristics of images, especially luminance (Harding & Bloj, 2010) attract our attention, as do surprising elements in an image (Itti & Baldi, 2009). However, to understand how humans interact with the myriad digital images of unknown veracity presented to us each day, it is also of interest to understand to what extent we can discern 'faked' photos, what our attitudes are towards them, and how we interpret the meanings presented in them.

The research and experiments undertaken in this research focused on a sequence of discovery with a strong element of human eye gaze tracking while viewing manipulated and unmanipulated images. The sequence begins with understanding people's attitudes towards manipulated images through a survey, continues through a series of eye gaze tracking and verbal questioning experiments designed to understand how people see manipulations in images, then tests whether classifiable patterns exist in the way we use our gaze to discern manipulations, and finishes with designing an image authentication framework that responds to what was learned.

Table 3-1 on the following page maps the research activities against the research goals identified in the Introductory summary and scope section 1.4.

Table 3-1: Map of research activities and goals

	Snap Survey	Experiment A	Experiment B	ANN analysis	Framework Design Spec
Research activity type	Survey	Eye gaze / verbal questioning using elements of .msci file format (n=12)	Eye gaze / grounded theory with advance training in types of image manipulation (n=80)	Analyse data with artificial neural net (ANN)	Draft design specification
Relevant paper		"Comparing eye gaze tracking to reported perceptions of manipulated / unmanipulated images"	"Imperfect understandings: a grounded theory and eye gaze investigation of human perceptions of manipulated and unmanipulated digital images"		Appendix B: Draft Design Specification
Research goals					
Understand the attitudes towards / significance of the problem of image manipulation to people	✓	✓			
Determine how people interpret manipulated images		✓	✓		
Determine if people can understand image manipulations better when images are presented in a trial file format that presents manipulated images together with their original version		✓			
Learn if people can be trained in discriminating manipulated images			✓		
Learn if computers can be trained to discriminate between eye gaze evidence				✓	
Determine what people need in an end-user solution to allow them to understand credibility of images	✓	✓	✓		
Design an image authentication framework that responds to known issues and this research	✓	✓	✓	✓	✓

3.1 Snap Survey: Attitudes toward image tampering

To obtain a broad perspective on people's attitudes toward image manipulation, I employed a survey that could be filled in within a few minutes. The survey methodology uses a Likert scale of 1-10 to determine participants' attitudes toward 4 questions:

- How easy is it for you to tell if a digital photograph is manipulated?
- How much do you care if the digital photograph you are looking at has been manipulated?
- How much do you think it matters in society if many digital photographs are manipulated?
- How concerned are you that there is currently no authentication process for digital photographs?

3.1.1 Methodology

The survey was given opportunistically between 2012 and 2016 at the events shown in Table 3-2.

Table 3-2: Snap survey events

Group	Date and location	Event	n =107	Subjects
1	April 2012 - ANU Centre for Higher Learning & Teaching	5-minute seminar presentation Graduate Teaching Program	5	Fellow participants in a multi-week training program presenting on our diverse areas of research
2	May-June 2013 - ANU Medical Interaction Lab	Experiment A	12	Largely first year university students (web development and design course)
3	October 2012 - ANU School of Art especially Dept of Photography	Presentation to School of Art "Work in Progress Conversations"	12	Research students and staff, many of whom were photographers
4	February 2014 - ANU Research School of Computer Science	Early Career Academic Cross-lab visit	6	Fellow early career academics
5	July 2015 - Barcelona, Spain	Paper given at conference - multimedia and human computer interaction	12	Conference attendees (Multimedia and Human-Computer Interaction)
6	April 2016 - ANU CECS April 2016	Girls in ICT Day	60	Girls in years 9-12 with an interest in information and communication technology

The survey was normally given in two parts. Part 1 contained the 4 questions only and subjects were given 1-2 minutes to complete the questions. The surveys were picked up immediately afterward. Subjects were then exposed to a presentation on the subject of image manipulation, and following this, were given the same questions again in Part 2, with a request that they score the questions afresh and provide any comments. The ‘before’ and ‘after’ versions of the survey were then matched up using pre-prepared codes on the pages and the variations in stated attitudes were compared.

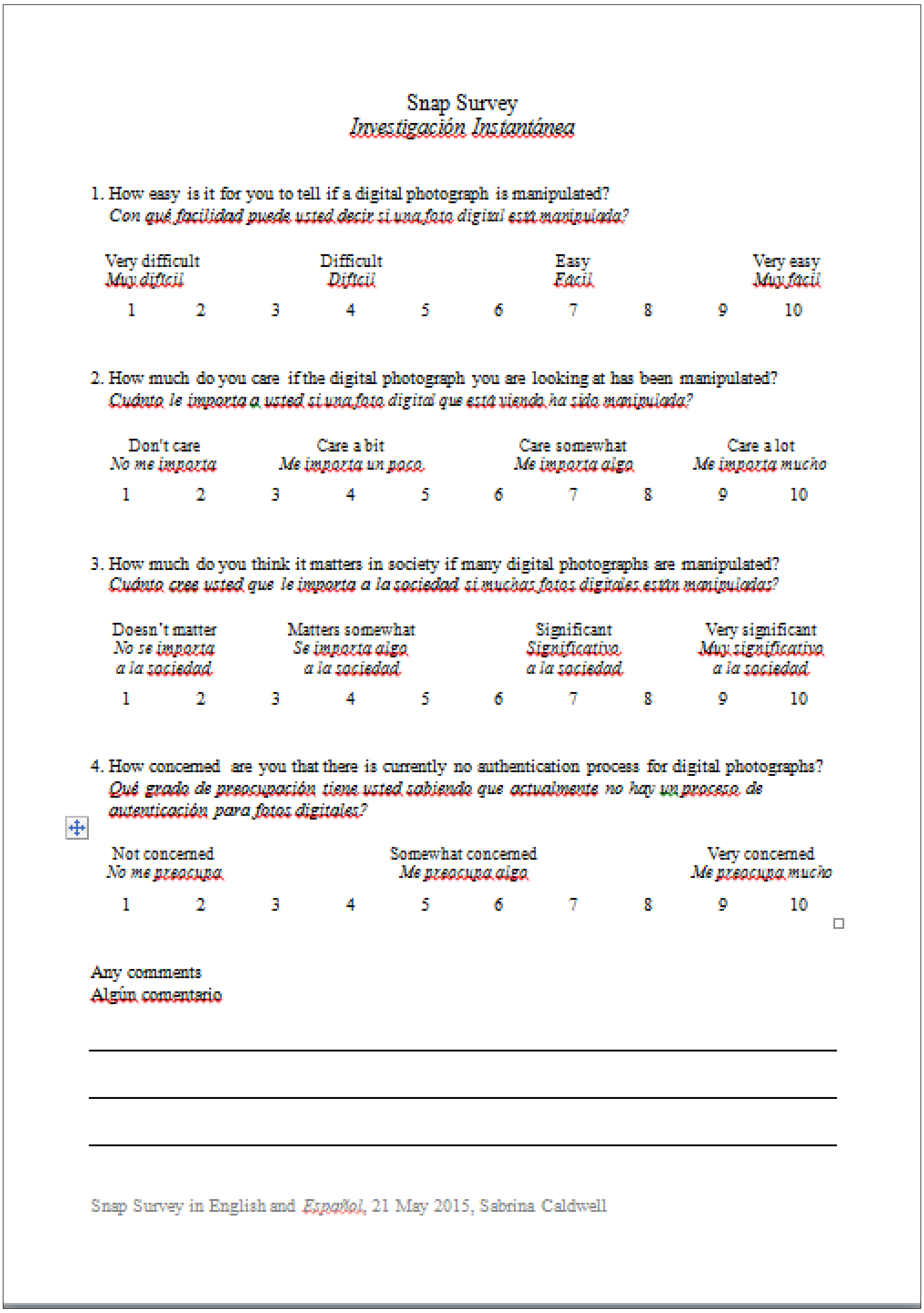
The version of the survey given in Barcelona was provided in both English and Spanish. Part 2 of the survey (which incorporates the same questions as part 1 with the sole addition of a comments section) is illustrated in Figure 3-1.

The first question differs from the remaining three in that it calls for respondents to self-assess their skill in spotting image manipulations. The remaining three questions required respondents to express their attitudes towards image tampering and the role it plays in society.

Responses were analysed on a group and collective basis, and Likert scores were analysed to assess:

- Initial attitudes of respondents
- Any change of attitudes by respondents after seeing a presentation on issues of manipulated images
- Differences and similarities between groups

Comments received provided additional insight into respondents’ attitudes and the Likert scores.



Snap Survey
Investigación Instantánea

1. How easy is it for you to tell if a digital photograph is manipulated?
Con qué facilidad puede usted decir si una foto digital está manipulada?

Very difficult *Muy difícil*			Difficult *Difícil*			Easy *Fácil*			Very easy *Muy fácil*
1	2	3	4	5	6	7	8	9	10

2. How much do you care if the digital photograph you are looking at has been manipulated?
Cuánto le importa a usted si una foto digital que está viendo ha sido manipulada?

Don't care *No me importa*			Care a bit *Me importa un poco*			Care somewhat *Me importa algo*		Care a lot *Me importa mucho*	
1	2	3	4	5	6	7	8	9	10

3. How much do you think it matters in society if many digital photographs are manipulated?
Cuánto cree usted que le importa a la sociedad si muchas fotos digitales están manipuladas?

Doesn't matter *No se importa a la sociedad*			Matters somewhat *Se importa algo a la sociedad*			Significant *Significativo a la sociedad*		Very significant *Muy significativo a la sociedad*	
1	2	3	4	5	6	7	8	9	10

4. How concerned are you that there is currently no authentication process for digital photographs?
Qué grado de preocupación tiene usted sabiendo que actualmente no hay un proceso de autenticación para fotos digitales?

Not concerned *No me preocupa*				Somewhat concerned *Me preocupa algo*				Very concerned *Me preocupa mucho*	
1	2	3	4	5	6	7	8	9	10

Any comments
Algún comentario

Snap Survey in English and Español, 21 May 2015, Sabrina Caldwell

Figure 3-1: Part 2 of the English-Spanish version of the Snap Survey on attitudes towards image manipulation

3.1.2 Results

It was clear from the outset that respondents were not indifferent to issues of image manipulation and the effect of manipulated images in society. Whether before or after discussions of image credibility issues, respondents expressed concern about the credibility of images and were not confident they could spot tampered images (Table 3-3).

Table 3-3: Mean values of all snap survey respondents

		Mean	SD	Likert representation
Q1	How easy is it for you to tell if a digital photograph is manipulated?	4.74	1.9	Difficult
Q2	How much do you care if the digital photograph you are looking at has been manipulated?	6.10	2.3	Care somewhat
Q3	How much do you think it matters in society if many digital photographs are manipulated?	6.58	2.5	Significant
Q4	How concerned are you that there is currently no authentication process for digital photographs?	6.15	2.4	Concerned somewhat

n=107

When taken as a whole, there was very little difference in participants' views after being exposed to issues of image manipulation in comparison to their views before (Table 3-4), that is to say they remained concerned and did not generally feel more confident they could distinguish manipulated images.

Table 3-4: Mean values 'before' and 'after' presentation

		Before		After		Diff	
		Mean	SD	Mean	SD	Mean	SD
Q1	How easy is it for you to tell if a digital photograph is manipulated?	5.14	1.7	4.26	1.9	-0.88	0.3
Q2	How much do you care if the digital photograph you are looking at has been manipulated?	6.05	2.2	6.16	2.4	0.12	0.3
Q3	How much do you think it matters in society if many digital photographs are manipulated?	6.58	2.5	6.51	2.7	-0.08	0.2
Q4	How concerned are you that there is currently no authentication process for digital photographs?	6.13	2.6	6.43	2.7	0.30	0.1

n=107

However, responses to the survey tended to be group dependent; different groups responded differently to the questions. Most notably, views from a group largely composed of photographers from the ANU School of Art, ran strongly against the trend. This can be seen in the red band of the graph below (Figure 3-2), in which this group demonstrated less concern about manipulated image issues (Q2-Q4). The group of year 9-12 girls were also somewhat less concerned, though they also expressed the greatest difference of all groups in their confidence in identifying images that had been manipulated, with their belief in their ability reduced by 30%.

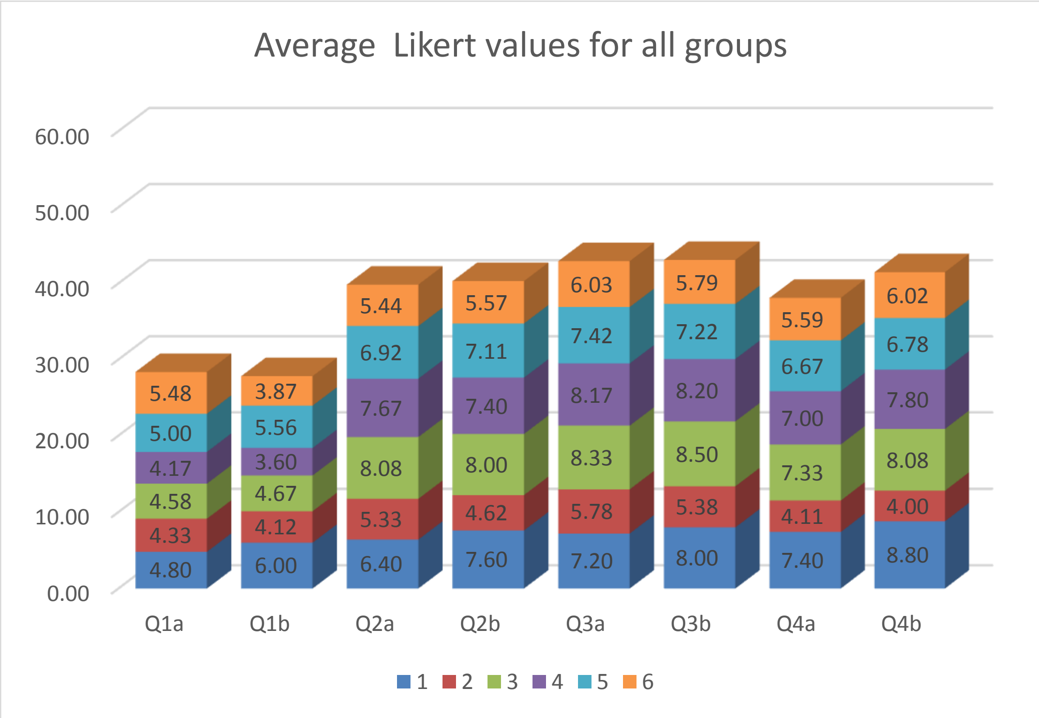


Figure 3-2 Snap Survey responses by group and question before exposure to discussion or presentation of manipulated image issues (a) and after (b). Groups are attendees 1) Graduate teaching program, 2) ANU School of Art, 3) Experiment A, 4) Early Career Academics cross-lab visit, 5) MHCI'15 conference, 6) Girls in ICT Day.

On average, participants self-assessment of their ability to spot fake photos decreased 17% from their before to after assessments (from 5.14 to 4.26 with lower number denoting increased difficulty) indicating that they had downgraded their ability, and now believed it was more difficult to spot manipulations).

Of the comments received, the most prevalent idea expressed was one of context. Participants noted the multi-faceted role of photographs:

> *The point about the division between "art" and "reality" is a very pertinent one – especially in the realm of propaganda and historical accuracy.*

– Early Career Researcher, 2014

> *Depends on the context (referred to Q1-3)*
>
> – Graduate teaching program participant, 2012

> *Manipulation only matters in a forensic context (legal, scientific, etc.) or represented in public media as accurate reportage.*
>
> – Dept of Photography respondent, 2012

> *Issues: 1) it's not whether an image has been manipulated, but how and why; 2) disclosure may be obligatory; 3) purpose of image maker or distributor; 4) mis-representation – what does the image purport to provide evidence in support of?*
>
> – Dept of Photography respondent, 2012

Another frequently expressed idea was doubt that photographs are capable of representing reality:

> *Isn't "manipulation" difficult to define? Even adjusting exposure on a camera is a form of manipulation. Choosing where to point a camera manipulates reality. Thus, I don't given any special status to photographs over other images (paintings, graphics); they are someone's representation of their perceived reality.*
>
> – Conference attendee, Barcelona, 2015

> *For me as a photographer there is no truth. Everything is a manipulation so I never think of anything being real or truth but a trace or haunting to an original.*
>
> – Dept of Photography respondent, 2012

There was only one comment that pointed out the risk to society of abandoning the photograph as a form of representational communication:

> *Significance for society [is that] people may <u>stop</u> believing in real photographs.*
>
> – Early Career Researcher, 2014

3.2 Experiments

As a step in further understanding aspects of this computer-mediated visual communication, I undertook experimentation using both eye gaze tracking and verbal questioning to compare what subjects see, as represented by their eye gaze tracking results, to what they perceive, as represented by grounded theory analysis of their verbal responses.

The first experiment resulted in the paper “Comparing eye gaze tracking to reported perceptions of manipulated and unmanipulated images” and the second experiment resulted in the paper “Grounded theory and eye gaze investigations in human perceptions of manipulated images”

These experiments had similarities but also differences. In both experiments important goals were to identify whether participants could identify manipulated images, and how they interpreted the meaning of those images. In the experiments, participants’ eye gaze was recorded using two Facelab 5.0.2 infra-red cameras and a single IR light emitter pod centrally located below the monitor displaying the images. Eyeworks v3.8 was used for experiment design, delivery, and aspects of analysis, and to record video evidence of each experiment.

In the experiments the verbal responses to each question were collected via video and audio recordings as well as notes. The questions included asking participants whether the images they viewed were manipulated, and in order to understand how they interpreted the images they viewed, participants were asked to describe their understanding of the meaning of the image.

However, experiment design and goals were different in each. In experiment A, the attitude snap survey described above in 3.1 was continued. Participants’ attitudes towards digital images were canvassed by integrating the snap survey into the experiment design. To determine the ability of participants to identify manipulations with and without help, images were presented in three formats (elements of the .msci file format which forms part of the image credibility framework proposed in this thesis): a) standalone, b) with the original, unmanipulated image for comparison, and c) standalone, original, and difference map images were presented.

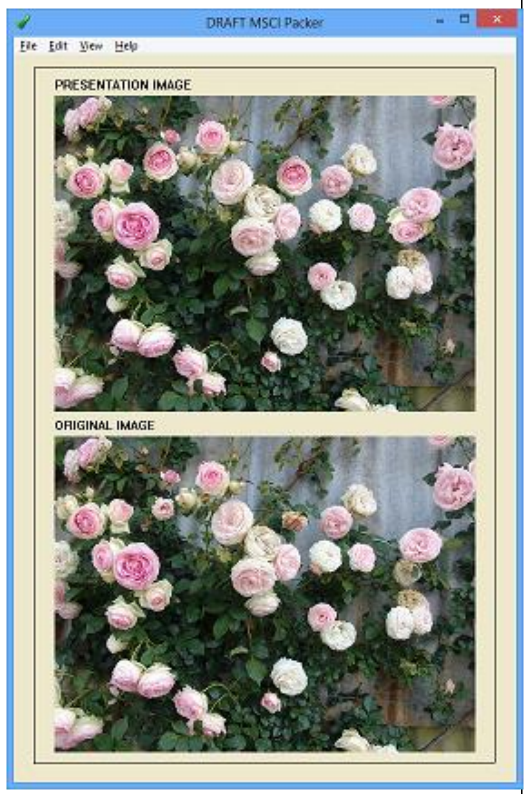
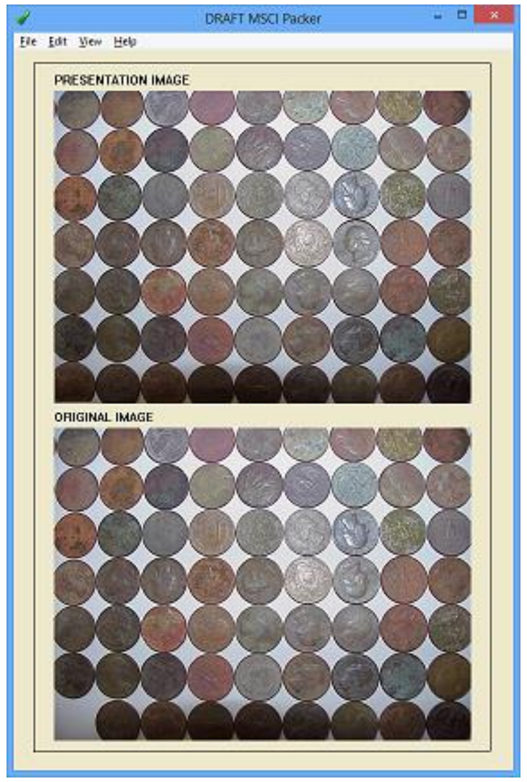
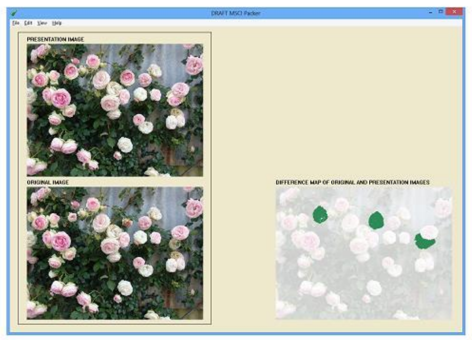
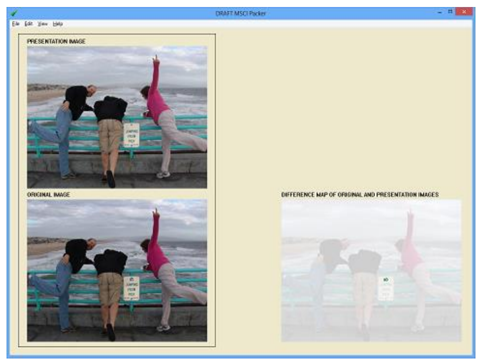
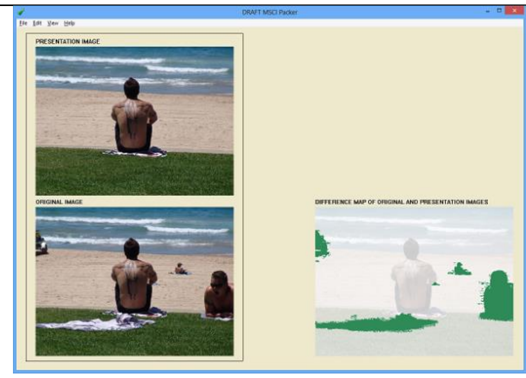
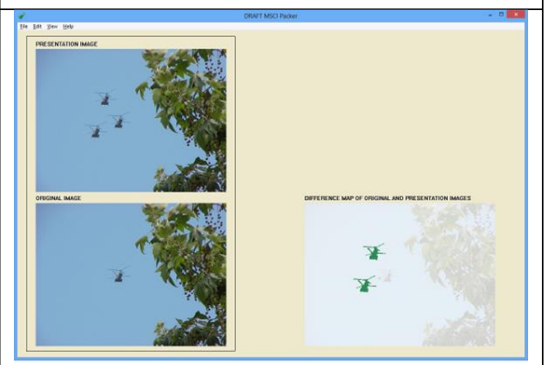
Set	Roses copy/move	Pier retouch	Coins copy/move	Man on beach retouch	Helicopter copy/move
Set 1- Standalone image			Not shown in set 1.	Not shown in set 1.	Not shown in set 1.
Set 2 – Comparison original					Not shown in set 2.
Set 3 – Comparison original and difference map					

Figure 3-3: Examples of images shown to subjects in Experiment A

This page intentionally left blank.

The following paper (Caldwell, Gedeon, Jones, & Henschke, 2015) outlining the results in detail was published in the Australian Journal of Intelligent Information Processing Systems. Vol 14 No 3. Note that the table / figure numbers and citations have been integrated with this thesis for ease of reference.

Comparing eye gaze tracking to reported perceptions of manipulated and unmanipulated digital images

Caldwell, S., Gedeon, T., Jones, R., Henschke, M.

Abstract

To investigate human perceptions of image manipulation at both the conscious and non-conscious levels, we compared participants' verbal reporting of image manipulation to data recordings of their eye movements while viewing 36 images of varying manipulation levels. To further understand subjects' ability to use image comparison tools to aid in manipulation detection, variants of a trial 'image packaging' software provided two levels of image comparison support tools.

Keywords: Eye gaze; manipulated images; detect image manipulation

3.2.1.1 Introduction

Increasingly, we encounter our information about the world in image form. At the same time, the ability of humans to manipulate images is greater than at any time previously in history. While there is research on the use of manipulated images in advertising[2] there is very little understanding of the effects of ubiquitous photo manipulation such as is rife in social media and family photos.

---

[2] For example effect of airbrushed models on teen body image (Grabe, Ward, & Hyde, 2008)

As a step in understanding this phenomenon, this experiment uses both eye gaze tracking and verbal questioning to compare what subjects see (as represented by their eye gaze tracking results) and what they perceive (as represented by their question responses) when provided with both standalone images and images that have been packaged with additional assistive information.

This experiment investigates peoples' ability to see manipulations in images, and seeks to identify whether providing additional comparison images along with the presentation image enables participants to identify manipulations in images more accurately and/or quickly.

In addition, the experiment attempts to determine how subjects interpret images in relation to any manipulations they contain.

3.2.1.2 Materials and Methods

Twelve volunteer participants undertook eye gaze tracking and verbal questioning as they viewed images of photographs ranging from unmanipulated to strongly manipulated. The participants' mean age was 32.7 (SE 11.1) years.

Facelab 5.0.2 by Seeing Machines (Seeing Machines, 2016) was used to track eye gaze with two infra-red (IR) cameras and a single IR light emitter pod centrally located in front and below the monitor displaying the images. Eyeworks v3.8, also by Seeing Machines, was used for experiment delivery, recording and analysis.

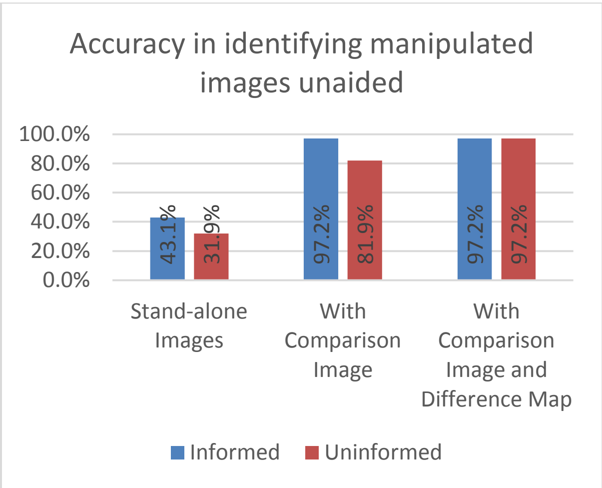


Figure 3-4 Image manipulation identification accuracy

Subjects were shown three sets of 12 images each comprising 3 unmanipulated images and 9 images manipulated by splicing in or erasing elements of varying sizes from the scene. The first set comprised standalone images. The second and third sets were presented in a mobile, self-contained image format (MSCI), a trial 'image packaging' software currently under development. The second set presented bundled images in which a presentation image was accompanied by the original image for comparison. This configuration was repeated in the third set but also accompanied by a 'differences map' image that highlighted any changed pixels.

The images were chosen to cover a wide variety of topics reflective of the types of images we encounter in everyday life through news, social media and email. They included family, military, nature, politics, landscapes, and society. It would not be possible to cover every type of photograph and manipulation in a reasonably sized experiment, and so representative selections were made. These included

Images within sets were varied in order using Latin square randomisation to avoid any ordering bias. In each subsequent set some images were repeated in the new format to identify how much assistance subjects needed to identify manipulations.

Subjects were assigned to one of two sub-groups, those who had pre-existing familiarity with image manipulation issues through exposure to the authors' research (informed), and those who did not (uninformed). In some cases (4 subjects) an additional set of eight images were employed after the common part of the experiment, to further test whether focussed exposure to image manipulation predisposes subjects to identify manipulated images with increased accuracy.

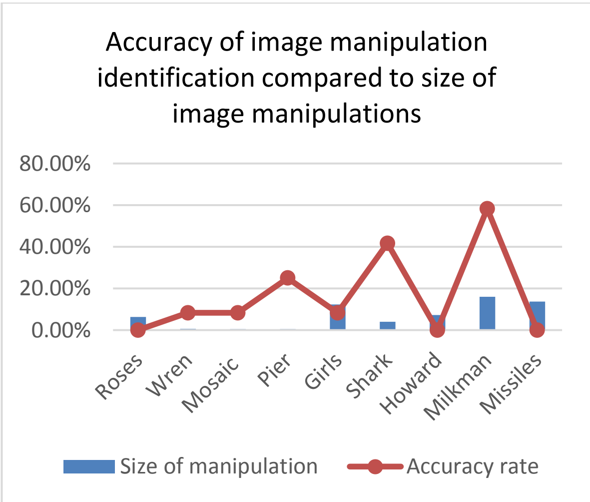


Figure 3-5 Image manipulation size to identification success

At the same time as their eye gaze was being tracked, participants were asked a short set of questions relevant to each phase of the experiment and their responses recorded. These questions targeted their perception of any manipulations that might appear in the images they viewed, and their interpretations of the images.

Finally, participants completed a short survey and responded to open-ended questions about their attitudes towards image manipulation.

3.2.1.3 Results

Overall, the ability of participants to verbally identify manipulated images with no assistance was weak, only 37.5% accuracy on average (Figure 3-4). This was despite eye gaze data indicating that subjects had looked directly at the manipulated areas with greater intensity than would be predicted by the area of the manipulated regions of the image (Table 3-5). Informed subjects were more likely to report image manipulations than uniformed subjects (43.1% vs 31.9%).

When given a comparison original image, participants' success rate at identifying manipulations more than doubled to 85.4%, although often they either could not say what had been changed or else misreported or under reported what had been changed. Eye gaze data indicated that participants gaze traversed the two images to identify and compare areas of difference. Again, there was a difference in accuracy between informed and uninformed subjects (97.2% vs 81.9%).

When also given a difference map in set 3 of the experiment, participants' success rate in identifying manipulated images increased to 97.2% for both groups, and the difference in quality of perception of manipulations increased. Eye gaze data indicated that participants' gaze referred to the difference map as an aid in locating manipulated areas presentation (manipulated) images as compared with the original.

Table 3-5 Comparing ratio of participant views to area size of manipulated regions

Regions of manipulation: % of participant views in relation to area of manipulated regions in Set 1									
	Roses	Wren	Mosaic	Pier	Girls	Shark	Howard	Milkman	Missile
% of gaze used in manipulated region(s)	12.7%	4.4%	5.7%	5.7%	15.6%	26.0%	27.2%	72.1%	36.7%
Area of manipulated region(s) in pixels	37863	4247	3079	2958	73494	6534	43617	95715	12665
Area of image in pixels	595337	594520	597908	593292	593025	159944	595856	599404	91921
Manipulated region % area of image	6.4%	0.7%	0.5%	0.5%	12.4%	4.1%	7.3%	16.0%	13.8%
% of views in relation to % of area	200%	611%	1116%	1151%	125%	636%	372%	451%	266%

Using the difference map, participants could identify 90.1% of specific manipulations in detail. When offered a comparison map but not a difference map, participants accurately identified the specific manipulations only 68.1% of the time. Further, incorrect alternative explanations for the effects of manipulations were given, for example the insertion of three additional roses into an image was described by three participants as increased colour contrast.

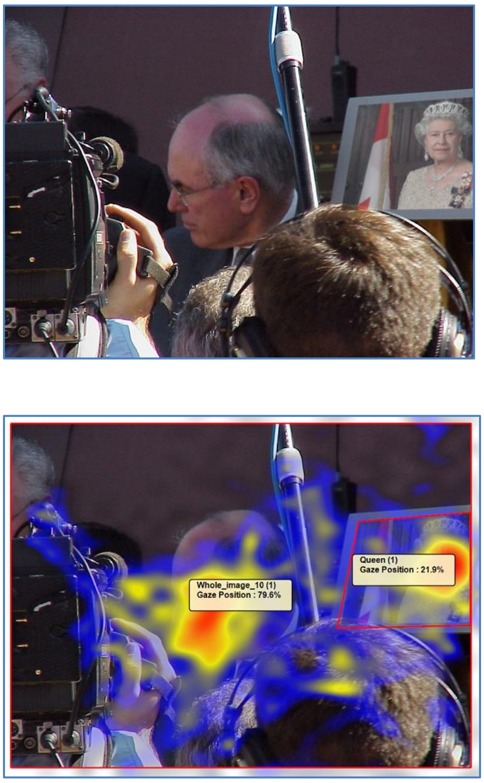


Figure 3-6: Set 1 standalone image with and without eye gaze 'heat map' annotation

Moving from the second set to the third set also increased speed of identification of manipulated images, with the time from image appearance to the decision point reducing from 12.4 to 5.3 seconds (standard deviations of mean in seconds were 5.16 and 2.99 respectively). In both cases this compares favourably with the time required for potential identification of manipulated images in Set 1 in which participants were presented with standalone images, 45.7 seconds (SD 7.1).

There was little correlation between the size of the manipulation and the accuracy rate (Figure 3-5).

Comparing eye gaze tracking to question responses indicated that subjects' eye gaze fixated on regions of manipulation in images even when they did not report the image as altered. Overall, while participants only identified 37.5% of the nine manipulated images in Set 1 (standalone images), their eye gaze rested in the regions of manipulation up to 11 times as often as the area the manipulations occupied would predict (Table 3-5).

As an example, the image of Australian Prime Minister (1996-2007) John Howard with a spliced image of Queen Elizabeth II from Set 1 standalone images was not verbally reported as manipulated by any participant despite the eye gaze 'heat map' demonstrating that the region of the spliced image was viewed intensively at an average 27% of eye gaze within area of the photograph and over 4 times more frequently than would be seen by area alone (Figure 3-6).

In responding to the survey questions, all participants stated that they cared about photo credibility (average 8.7 out of 10 indicating that on average all the subjects care a lot), societal implications of photo manipulation (average 9.1 out of 10 indicating this significantly matters), and the lack of photo authentication solutions available (average 8.25 out of 10 indicating concerned to very concerned).

Responses to the question “How easy is it for you to tell if a digital photograph you are looking at has been manipulated?” yielded answers clustered around the midway point (4.9 indicating a middle ground between difficult and easy).

In post experiment open-ended questioning, 10 (83%) of participants responded that they use one or more photo editing software packages including Adobe Photoshop, Microsoft Paint, Fireworks, Instagram, Gimp and Adobe Illustrator. The uses to which these software systems are put include cropping, red-eye reduction, colour adjustment, light adjustment, ‘fun filters’ in the case of Instagram, and making collages.

In response to the question of how they identified manipulations when viewing photographs, the strategies identified by participants were “searching for sharp edges,” “when things don’t look right, like one face on another person’s body,” “lighting effects,” “if the dimensions are wrong,” if things are too perfect,” and “shadows going the wrong way.” Some participants (25%) stated they did not know how to identify manipulations in images.

When asked what they look for in a photograph, with suggestions of aesthetics, meaning, or representations of reality, participants responded overall that they looked first for aesthetics, then meaning, then representations of reality.

3.2.1.4 Discussion

Participants brought a diverse range of understanding of photographic images to the viewing exercise. This understanding often informed their detection of image manipulation. Most significantly, participants who were aware that the research involved photo credibility were more successful in identifying manipulations (Figure 1), which may have resulted from them viewing the photographs with a more critical eye (that is, perhaps they were more likely to consciously pay attention to the results of their non-conscious identification of changed regions of the photograph).

Knowledge and life experiences generally played a role in participants understanding the meaning of the photographs they viewed. For example, of the 12 participants, only 1 articulated the connection between the image of Queen Elizabeth spliced into the photograph of then Prime Minister John Howard in Figure 3 above (John Howard's well-known monarchist views on whether Australia should be a Republic). Three subjects were unable to identify John Howard at all and focussed on the Queen or the media aspects of the image.

It was expected that manipulations of larger sizes would be spotted more readily than manipulations of smaller sizes, this was not the case. There was no obvious correlation between the size of the manipulation and the accuracy rate (Figure 2). In an image used in all three sets, in which it appears three people are about to jump over a pier rail next to a sign that reads "JUMPING FROM PIER" (from which the word 'NO' was erased from the image) the overall size of the manipulation was only 0.5% yet the accuracy rating of participants identifying the manipulation was 25%. At the same time, an image of missiles in which 13% of the image was an additional spliced image (The Telegraph UK, 2016) had a 0% success rate of manipulation identification.

Given the tendency of participants to rationalise elements in images (discussed below), it may be that a more defining characteristic of more easily discernible manipulations is

Figure 3-7: Hoax photo of cow on bonnet of BMW sedan

their saliency, i.e. the extent to which the elements added to or removed from an image contribute to the understanding of that image. This suggests that further experimentation to tease out the differences between apparent size and saliency impact may be useful.

While all participants stated at a level of 8.25 or over out of 10 that they cared about photo credibility, societal implications of photo manipulation, and the lack of photo authentication solutions, their verbal exposition when discussing the meaning of the photographs presented to them indicated that they were more likely to justify the oddness of the image than to question it.

This was true even when the subject of image manipulation had been discussed moments earlier. Short-term increased awareness of image manipulation issues (as represented by participation in the experimental study) appeared to have little effect in conditioning participants to look at photographs critically. Four participants who had been assigned to the untrained group prior to the experiment participation were also asked to view 8 additional images, 2 of which were unmanipulated and 6 of which were manipulated similarly to the common part of the experiment. Three responded "no" in response to the query "Do you believe that any of these eight photos has been

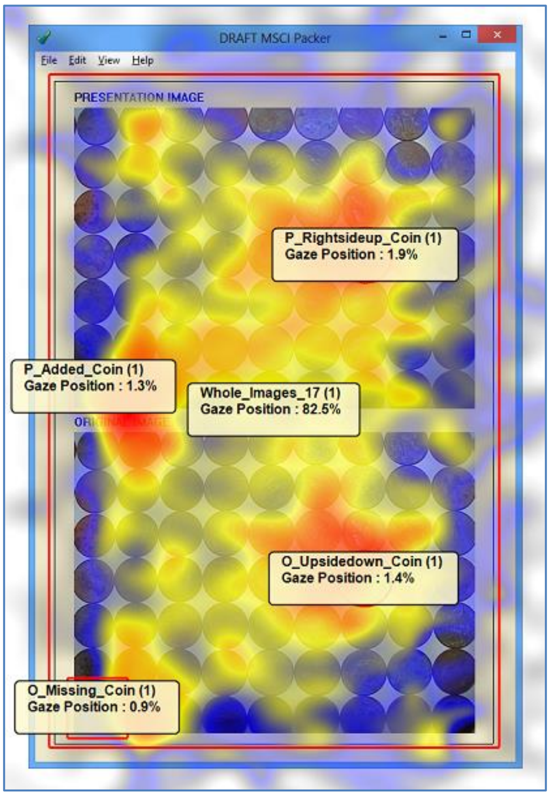


Figure 3-8: Manipulated coins photo: presentation image provided with comparison original

faked?" The 4th subject responded yes but could not identify more than one faked photograph from the 6 presented.

For example, although common sense would dictate that the photograph in Figure 3-7 had been manipulated, participants explained it away with justifications such as "maybe the car was warmer to sleep on than the snow," "relates to the use of leather in cars," or an inspirational message of unknown origin "don't think that anything is impossible."

This matches the uncertainty of responses to the survey question about ways to identify photograph manipulations, as well the verbal exposition and eye gaze data in which participants often used words indicating uneasiness with a photograph as they looked at manipulated elements in images, such as one participant commenting "that shouldn't be up there" as her eye gaze rested on the cow in the image at Figure 3-7.

It was noted that a 'hiding' effect occurred when additional, less obvious alterations were included in an image. In Figure 5 an image of a field of coins was presented to subjects for the first time in Set 2 of the experiment wherein subjects were offered both a presentation image and the original for comparison. Although most (11 of 12) noted that a coin had been added to the original, only one participant noted that another coin in the image had been rotated 180 degrees. This was despite eye gaze tracking identifying that subjects looked at the rotated coin more (2.3% of area within photo) than the added coin (1.6% within photo).

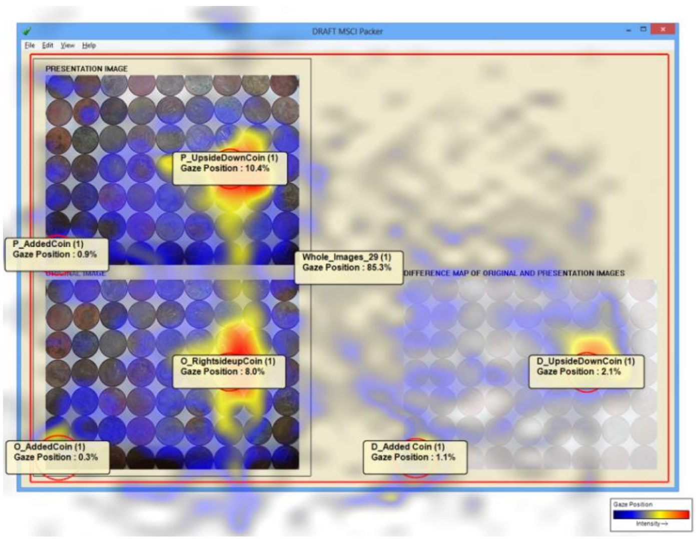


Figure 3-9: Subjects seek second manipulation when offered a difference map

When the coins photograph was presented again in the third phase of the experiment in which they were offered a difference map identifying changed pixels (Figure 3-9), all subjects used an eye gaze strategy that compared the pixels demarcated in the difference map as changed to find and identify the rotated coin in the presentation image by comparison with the original image.

In some cases the use of the difference map in Set 3 enabled participants to confirm previously identified differences noted in Set 2, somewhat like an answer key. In other cases the difference map provided participants with information that enabled them to identify that there were differences or additional differences in images they had previously passed as unchanged. Overall, using the image configuration in Set 3 provided the greatest level of accuracy (97.2% manipulated images identified, 90.1% of all manipulated regions identified) and speed (5.3 seconds on average).

It may be that one reason subjects are more likely to explain than to question manipulated images is that they want to believe they can spot fakes and therefore seek alternative explanations for unlikely elements in the images. Conversely, it may also be the case that subjects feel they can't spot fakes and therefore focus on and rationalise the meanings of the images.

Figure 3-10: Manipulated sea anemones in frozen pond photograph

These rationalisations can be quite unexpected, as in the case of a participant attempting to explain the juxtaposition of a trio of sea anemones and a frozen pond (Figure 3-10). In this case the participant focussed on the anemones being sea creatures and reinterpreted the snow on the edging stones as salt.

It is worth noting that the co-existence of high levels of concern about photo manipulation and the tendency to justify rather than identify manipulated images is dichotomous. It may be that photographs have not yet shed their cachet of being representations of reality; subjects are conditioned to look at images as 'real'. This would also be useful to examine in future experiments.

3.2.1.5 Summary

Comparing eye gaze tracking to question responses reveals that subjects may see more of the changes in manipulated images than they consciously report.

It is not necessarily the case that larger manipulations are more easily seen. The saliency of the manipulation may influence the identification success rate, and this bears further investigation.

It appears that it may be that when an image has an obvious manipulation, other lower profile manipulations may not be consciously identified even when they are viewed by the eye gaze.

Subjects who were aware that this research involves manipulated images (the informed group) looked more critically for possible manipulations and performed better when presented with both standalone image and comparison images than those who were uninformed.

Exposing participants to additional standalone images subsequent to the experiment proper did not result in improved performance comparable to the 'informed' cohort in identifying manipulations.

There is a dichotomy between the high levels of concern expressed about photo manipulation and the tendency of participants to explain away manipulations instead of identifying them.

The two levels of MSCI image bundling assisted subjects in perceiving image manipulations more accurately and quickly.

- end of paper -

3.2.2 Experiment B Part 1

This experiment used verbal questioning to identify what participants perceived (as represented by their question responses) when viewing both manipulated and unmanipulated images. In addition, eye gaze tracking was used to a) collect data about what participants physically see (as represented by their eye gaze tracking results) and b) to determine how that relates to what they report that they see. The experiment also attempted to determine how participants interpret images in relation to any manipulations they contain.

Sixty volunteer participants undertook eye gaze tracking and verbal questioning as they viewed manipulated and unmanipulated photographic images.

For all images, each participant was asked the following questions in order and their freeform answers were recorded:

- What in this image do you find interesting; what attracts your attention?
- Do you believe this image has been manipulated or has not been manipulated? Why?

In two selected cases (images 8 and 13, Figure 1) participants were told the nature of the manipulation (spliced missile and spliced woman respectively) and asked if they could identify which had been added.

A significant difference in this experiment to Experiment A is that participants were explicitly trained in issues of image manipulation: participants read and completed a tutorial on the topic of "*Digital Images*". This was to pre-familiarise participants with current trends in image processing and credibility and to give them the maximum chance of successfully identifying image tampering. This is preferred to naïve viewing due to an expectation that with increased use of images in everyday life a higher level of visual literacy can be expected to develop over time. The tutorial was taken from the first year ANU computer science course Comp1710 – Web Development and Design. The tutorial was composed of 3 topics: general image tampering, provenance issues of copyright and intellectual property, and image credibility.

The following paragraphs are part of the text given to subjects.

> Image tampering has been around since the advent of the photographic process. In the 150 years commencing with the invention of conventional photography by Henry Fox

> Talbot and Louis Jacques Mande Daguerre, and before digital photography was introduced, photographers were staging images and/or creating seemingly real but actually false photographic prints or photo art pieces crafted from disparate negatives in photographic darkrooms.
>
> Digital photography and the long arm of the Internet increased the problem of photo manipulation. Image manipulation software has become inculcated into photographer's postprocessing of photographs, offering easy access to an extensive palette of image tampering tools. Such manipulated images are now common; although usually manipulated for fun or art some photos are manipulated for political or commercial ends. Presently, photos are often illusive electronic constructs, globally distributed at the speed of light with little context or explanation.
>
> Despite attempts by some camera manufacturers, no authentication process has yet been successfully implemented. However, a range of digital image forgery detection techniques have been developed in recent years. Collectively these techniques are known as digital image forensics, a field that analyses images to determine image veracity through identifying image manipulation artifacts. Interest in and development of digital image forensics techniques is increasing due to the potential of image manipulation to impact on medicine, justice, news reporting and the legal and accounting professions. Recently, defendants have been successful in rejecting photographic evidence based on the fact that they cannot be authenticated.

The subjects' assimilation of the information was tested with questions like the one provided below.

> Q: Manipulating photographs have implications for law enforcement and the justice system in that:
>
> Select one:
> a. Defendants don't like to see photographs of themselves.
> b. There are too many digital photographs to sift through when an incident happens.
> c. The resolution of digital photographs is not as good as conventional photographs.
> d. Because people are innocent until proved guilty, the onus is on the prosecution to prove that the photos they use are unmanipulated and there is currently no authentication process to do that.
> (Answer = d)

Manipulations were limited to *copy/move* changes in which one part of an image was cloned and repositioned within the photograph and *splicing*, in which some or all of a secondary image was incorporated into the primary image. One additional image was globally changed from colour to black and white (Fig 3); both images were presented. It was not possible to represent every type of photograph within the confines of a single experiment. However, the general cohort of images was chosen to (to the extent possible) the types of images to which participants could be expected to be exposed on a daily basis. The images were chosen to cover a wide variety of topics, including family, military, nature, politics, landscapes, and society. Specifically:

Figure 3-11: – Juvenile chameleon

Figure 3-12: Party-goers, front woman spliced

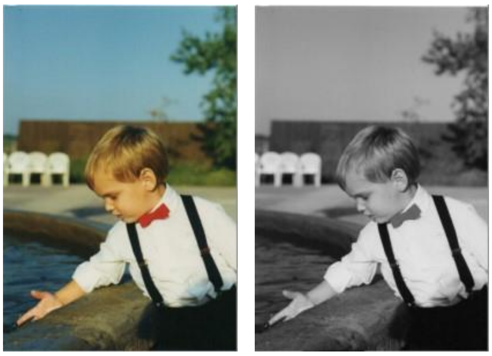


Figure 3-13: Boy (colour and b&w)

Figure 3-14: Car cow

Image 1 – Chameleon on a matchstick – This photograph is unmanipulated, but was chosen as being likely to be mistaken for manipulated due to its surprising nature.

Images 2 and 3 – ‘Bobby’ – The first of this pair (Image 2) is unmanipulated; the second (Image 3) has been manipulated by changing the colours to black and white. The purpose of the pair was to elicit opinions from subjects as to which is the unmanipulated photograph.

Image 4 – “Holly” is photographed at low (<1Mb) resolution. Cats are a common image type consumed online, and it was of interest to see if low resolution impacted on subject perceptions of the photograph’s credibility.

Image 5 – “Bailey” is again a photograph of the ubiquitous subject of cats but in this instance, was photographed at much higher resolution (8Mb).

Image 6 – “Fading Away” by Henry Peach Robinson is a historical image known to be extensively manipulated by the photographer and was included as a counterpoint to the more modern photographs as one subjects may encounter in an article or similar text.

Image 7 – “Zebra in a lingerie shop” is an unmanipulated photograph, but like Image 1 could potentially be mistaken for a manipulated photograph due to its surprising nature.

Image 8 – “Missiles” is a photograph representative of news photography. It is known to be manipulated (third missile cloned in).

Image 9 – “Cow on car” is a manipulated photograph credited to the Surrey Hills police, meant to be humorous, and therefore representative of entertaining images that circulate via social media.

Image 10 a and b – “Coins” was chosen for its geometric symmetry which emphasises a gap in the original (10a) which might suggest a manipulation that did not exist. The gap is filled in the manipulated version (10b).

Image 11 a and b – “John Howard” is another political image. In 11a there is a light reflection screen that in the manipulated version 11b has been altered to show an image of the Queen of England (then Prime Minister John Howard was known to be a monarchist).

Image 12 a and b – “Anemone” is a straightforward natural scene in 12a, but a grossly manipulated image in 12b, meant to provide an opportunity to see distinct differences in the eye gaze and verbal responses of the two cohorts.

Image 13 a and b – “Girls” is another representation of a common social media type image. In 12b an additional girl has been added.

Image 14 a and b – “Pier” is a photograph that might be found in any family photo album. The manipulation in 14b is very small but alters the meaning of the image: the sign saying “NO JUMPING FROM PIER” has been altered to read “JUMPING FROM PIER.”

Sixty (44 male, 16 female) participants with an average age of 24.3 years (SD 8.9, range 18 to 60) took part in the study. Participants were sourced largely from a first year computer science course (COMP1710) offered at the Australian National University (ANU). Of the total participants 41 were COMP1710 students. The remaining participants had a mixed demographic.

Facelab 5.0.2 by Seeing Machines was used to track eye gaze with two infra-red (IR) cameras and a single IR light emitter pod centrally located in front and below the monitor displaying the images. Eyeworks v3.8, also by Seeing Machines, was used for experiment delivery, recording and analysis.

Participants were assigned to one of two cohorts of equal size (30 in each cohort). Each cohort was shown 14 images comprising a mix of manipulated and unmanipulated images. Five of the images were shown in different forms to the two cohorts, with Cohort 1 viewing the images in their manipulated form, and Cohort 2 viewing the images in their original, unmanipulated form. This was done to tease out differential eye gaze indicating visual attention to salient and/or manipulated regions of images. Benchmarks for identifying accuracy in identifying unmanipulated and manipulated images were established by calculating the results of the 9 images common to both cohorts.

As described above in this section 3.2.2, both cohorts were familiarised with principles of digital photograph manipulation through an online textual explanation of manipulation techniques including copy/move, splicing, and global transformation such as used in the experiment.

At the same time as their eye gaze was being tracked, participants were asked a short set of questions relevant to each image and their responses recorded. These questions targeted their perception of salient elements and manipulations that might appear in the images they viewed, and their interpretations of the images.

Eye gaze fixations were compared in aggregate to verbal responses to determine whether visual attention was paid to areas of image manipulation regardless of verbal confirmation. To gain an understanding of the participants' conscious perception of what interested them in each image, they were asked the question "What in this image interests you; what draws your eye?" To directly test their conscious ability to perceive manipulations in images, participants were asked if the image had been manipulated. In this instance, manipulation was described as whether they believed that the photograph had been subject to any changes from the original photograph as taken.

3.2.2.1 Results

The ability of participants as a whole to verbally identify whether images presented to them were manipulated or unmanipulated was poor, with an overall 64.2% mean success rate (SD ±15.4%) across a total of 629 image views.

To establish a benchmark for predicting accuracy across the two cohorts, the results from the 9 images common to both cohorts were calculated. Using the results of only those common images yielded benchmarks of 63.4% (SD 23.07) accuracy rate for unmanipulated images, a 66.7% (SD 14.51) accuracy rate for manipulated images, and an overall accuracy rate of 65.0% (SD 18.61). The broad range of the standard deviations for the benchmark figures reflects the nature of the images in which some images were quite difficult to determine manipulation, while others were easier. In some cases the images were startling (for example Figure 3-11 and Figure 3-14) that evoke targeted fixations due to incongruity (Loftus, Loftus, & Mackworth, 1978), and in others, images were standard 'day-to-day' images such as might be found on Facebook (for example Figure 3-12 and Figure 3-13).

Results for both cohorts, and the differing outcomes for the cohort viewing an additional 5 manipulated images as compared to viewing the additional 5 images as unmanipulated originals were calculated. The results showed that both cohorts were well within the standard deviation of the benchmarks (Table 3-6). This suggests that there is little difference in success rates between identifying manipulated images as manipulated, and unmanipulated images as unmanipulated.

Table 3-6: Accuracy rates in verbal reporting of image manipulations

	Bench-mark	SD	Cohort 1	SD	Cohort 2	SD	Diff (Cohort 1)	Diff (Cohort 2)
% unmanipulated image accuracy results only	63.37%	23.07%	66.34%	19.80%	62.62%	23.14%		64.60%
% manipulated image accuracy results only	66.67%	14.51%	63.51%	20.06%	64.13%	13.34%	61.90%	
% overall accuracy	65.02%	18.61%	64.93%	19.27%	63.07%	20.30%		

3.2.2.2 Discussion

Despite pre-familiarisation with concepts of digital image manipulation, participants were poor at identifying whether images were manipulated or not. However, we note that the average accuracy rate of participants in this experiment (58.1%) compares favourably to the results in a previous experiment (Caldwell, et al 2015) in which participants with only a brief familiarisation with principles of image manipulation achieved a lower mean accuracy rate of 37.5%.

Comparing eye gaze tracking with question responses indicated that subjects' eye gaze fixated on regions of manipulation in images even when they did not report the image as altered.

The table in Table 3-7 shows some preliminary comparisons between the accuracy of participants identifying manipulations and whether they viewed the manipulated region. The figures were derived by comparing whether the participant verbally reported the image as manipulated to the recordings of their eye gaze fixations in the manipulated regions.

It is interesting to note that of the participants who accurately reported manipulations in images, more did not fixate on the key areas than those who did (88 'saw' the manipulations, 100 'did not see' them). This may relate to guessing and is part of continuing investigations of the data collection.

Table 3-7: Comparison of verbal reporting accuracy to eye gaze fixation in manipulated regions

	'Did see'	'Didn't see'
Correct	88	100
Incorrect	79	127

3.2.2.3 Summary

Despite pre-familiarisation with concepts of digital image manipulation, participants were poor in identifying whether images were manipulated or not.

Eye gaze tracking used to collect data about what participants physically see suggests that they may observe manipulations at a non-conscious level. The data collected from this experiment forms the dataset for further investigation into this observation in 3.3.

3.2.3 Experiment B Part 2 and "Grounded theory and eye gaze investigations in human perceptions of manipulated images"

Subsequent to Experiment B Part 1 and my initial findings, the experiment was expanded with the addition of 20 more subjects who undertook the experiment to add to the volume of data sets achieved, and in addition the data was subjected to new methods of analysis of different aspects of the data obtained. In particular, the qualitative data obtained from verbal questioning seemed to offer further insights into how we perceive manipulations in images, which was not picked up in quantitative analysis alone.

Accordingly, the new data set comprising the data from the initial 60 subjects and the added 20 subjects was analysed using Glaserian grounded theory (Glaser, 2007), a well-established qualitative analysis method used to interrogate verbal information and identify the concepts contained therein.

The resulting paper, which was awarded Best Paper at the Multimedia and Human-Computer Interaction conference in Barcelona, Spain in July 2015 is presented below.

The following paper (Caldwell, Gedeon, Jones, & Copeland, 2015) was published in Avestia: The Proceedings of the 3rd International Conference on Human-Computer Interaction, Barcelona, Spain. Note that the table / figure numbers and citations have been integrated with this thesis for ease of reference.

Proceedings of the 3rd International Conference on Human-Computer Interaction
Barcelona, Spain – July 13-14, 2015
Paper No. 308

Imperfect understandings: a grounded theory and eye gaze investigation of human perceptions of manipulated and unmanipulated digital images

Sabrina Caldwell, Tamás Gedeon, Richard Jones, Leana Copeland
The Australian National University, Research School of Computer Science
Acton, ACT, Australia
sabrina.caldwell@anu.edu.au, tom.gedeon@anu.edu.au,
richard.jones@anu.edu.au, leana.copeland@anu.edu.au

***Abstract* –**To investigate the extent to which humans are able to perceive manipulated images, and how they interpret these images, we use eye gaze tracking and grounded theory analysis of 80 participants viewing a combination of manipulated and unmanipulated images together with pre-familiarisation with image manipulation techniques. We find that accurate identification of manipulations is poor to moderate, and ability to identify *what* has been manipulated is poor. However, extended visual attention to manipulated regions was associated with greater accuracy. Grounded theory analysis shows that uncertainty in respect of veracity of images viewed is common across all images. Similarly, the use of logic influences participants' success rate in identifying manipulations.

***Keywords*:** Eye gaze tracking; manipulated images; detect image manipulation

1. Introduction

Increasingly, we encounter information about the world in image form (Kress, 1996). Users of Facebook alone upload more than 350 million photos every day (Facebook, 2013). At the same time, image manipulation is no longer restricted to specialist photographers, but is achievable by anyone with image editing software. Issues relating to inappropriate image manipulation have been noted in a range of disciplines including news reporting (Wheeler, 2002), medicine (Prasad, 2011) and scientific journals (Cromey, 2010).

There has been significant research on how we view images. Examinations of saliency in features of an image have shown that we focus on salient features of an image to discern its meaning (Itti & Koch, 2000), (Treisman & Gelade, 1980; Underwood & Foulsham, 2006), and investigations of eye gaze scan paths demonstrate that some characteristics of images, especially luminance (Harding & Bloj, 2010) attract our attention. However, to understand how humans interact with the myriad digital images of unknown veracity presented to us each day, it is also of interest to understand to what extent we can discern ‘faked’ photos, and how we interpret the meanings presented in them. In previous work the authors found participants’ success rate to be related to the availability of reference images (Caldwell, Gedeon, Jones, & Henschke, 2015).

As a step in further understanding aspects of computer-mediated visual communication, this experiment uses both eye gaze tracking and verbal questioning to compare what subjects see, as represented by their eye gaze tracking results, to what they perceive, as represented by grounded theory analysis of their verbal responses.

2. Materials and Methods

Eighty volunteer participants undertook eye gaze tracking and verbal questioning while viewing 14 manipulated and unmanipulated photographic images. Volunteers were sourced largely from a first year computer course and augmented with others from a range of sources. Participants’ mean age was 24.4 (SD 8.7); 53 were male and 27 were female.

Image manipulations used were *copy/move* (one part of an image cloned and repositioned within the photograph) and *splicing* (some or all of a secondary image incorporated into a primary image). An additional image was globally changed to grayscale from colour, with both images presented.

Participants were divided into two cohorts of 40 each. Both cohorts viewed 9 common images comprising 5 manipulated and 4 unmanipulated images. Each cohort viewed 5 additional images, differentiated so that one cohort saw the original unmanipulated versions of the images and the other saw the manipulated versions of these images. The images were chosen to cover a range of topics including social images, nature, and politics (Figure 3-15).

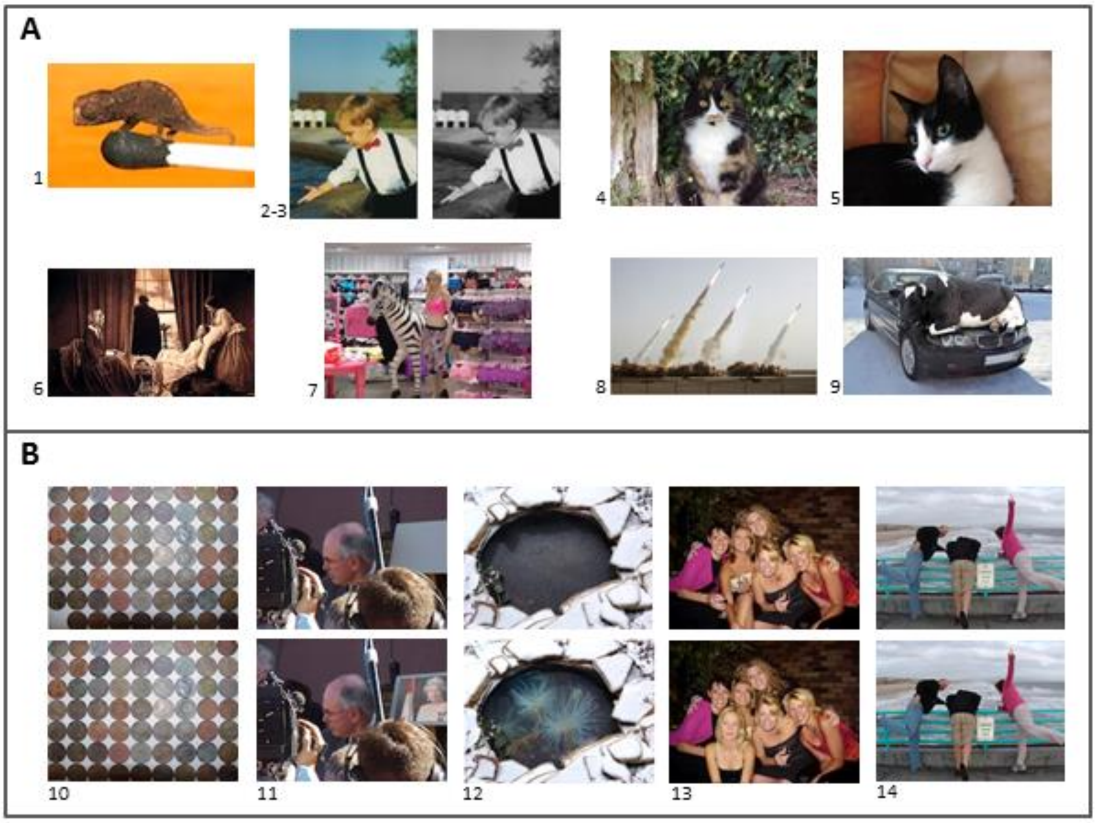


Figure 3-15- Common images (A) and images differentiated between cohorts (B)

To support their ability to perceive manipulated features in images, participants were presented with text describing image manipulation techniques prior to viewing images.

For all images, each participant was asked the following questions in order and their freeform answers were recorded:

- What in this image do you find interesting; what attracts your attention?
- Do you believe this image has been manipulated or has not been manipulated? Why?

In two selected cases (images 8 and 13, Figure 3-15) participants were told the nature of the manipulation (spliced missile and spliced woman respectively) and asked if they could identify which had been added.

The Glaserian grounded theory approach was used to analyse responses (Glaser, 2007).

Participants' eye gaze was recorded using two Facelab 5.0.2 infra-red cameras and a single IR light emitter pod centrally located below the monitor displaying the images. Eyeworks v3.8 was used for experiment design, delivery, recording and analysis, and to record video evidence of each experiment.

Regions of manipulation were transparently demarcated in relevant images and participants' eye gaze fixations were analysed to determine intensity of viewing of these regions.

3. Results

The overall ability of participants to state whether images presented to them were manipulated or not was poor to moderate, with a 56.0% mean accuracy rate based on 600 accurate identifications across a total of 1071 valid image views (reasons for invalidation included eye gaze registration failure and six participants stating they had previously seen an image). Participants had greater success correctly identifying images as unmanipulated (61.3%) than correctly identifying images as manipulated (50.1%).

Comparing these results to participants' eye gaze for the manipulated images (Table 3-8: Difference of participants' eye gaze viewing manipulated areas of images) shows that the relationship between eye gaze and accuracy varies widely between images with statistical significance of the difference ranging from not at all significant (0.941 for fixations and 0.908 for duration of gaze) to very significant (0.009 for fixations and 0.002 for duration of gaze). However, of the 8 manipulated images, 3 showed a significant difference between the mean fixations and durations of viewing the manipulated regions between those who stated that the images were manipulated and those who stated that they were not. Across all of the images, there was a significant difference ($p < .001$) between the 732 observations of fixation and duration of eye gaze of those correctly identifying images as manipulated and those who incorrectly identified them as unmanipulated. Fixations were: mean 38.4, sd 25.2, n=213 for those who stated image was manipulated, and: mean 27.3 sd 22.3, n=153 for those who stated image was unmanipulated; durations were mean 10.3 sd 7.6, n=213 for those who stated image was manipulated and mean 7.3 sd 6.6, n=153 for those who stated image was unmanipulated. When Images 10, 11, and 12 are omitted, the difference in fixations between these two outcomes remains statistically significant.

Table 3-8: Difference of participants' eye gaze viewing manipulated areas of images[3]

Image	n (total)	Stated: manipulated (correct)					Stated: unmanipulated (incorrect)					Sig (p<.05)	
		Mean Fixation	SD	Mean ∑ Duration (secs)	SD	n	Mean Fixation	SD	Mean ∑ Duration (secs)	SD	n	Fixation	Duration
3	63	63.4	24.43	15.39	8.29	42	56.7	25.64	15.14	8.32	21	0.329	0.908
8	67	27.0	18.07	7.66	6.08	39	28.6	23.10	8.11	6.43	28	0.765	0.772
9	67	30.2	20.72	9.60	7.56	54	29.0	17.42	8.76	6.45	13	0.837	0.689
10	20	9.3	4.35	7.75	6.68	7	3.1	2.42	2.32	2.49	13	**0.009**	**0.002**
11	37	38.5	19.09	8.34	4.77	21	25.1	12.03	5.45	3.67	16	**0.014**	**0.045**
12	35	54.5	18.82	13.02	6.65	25	32.1	15.21	6.91	5.11	10	**0.001**	**0.008**
13	34	30.1	19.32	8.24	5.97	13	30.6	18.17	8.69	5.92	21	0.941	0.836
14	36	17.0	9.40	7.54	7.93	12	12.3	6.64	4.28	2.79	24	0.143	0.192

The overall ability of participants able to identify *what* was manipulated was considerably lower, with only 132 correct identifications of what was manipulated out of 477 valid views of manipulated images, or 27.7%.

4. Discussion

Despite extensive pre-familiarisation with concepts of digital image manipulation, participants were poor to moderate at identifying whether images were manipulated or

[3] Note that manipulated Image 6 was omitted from Table 1 because although it is known that the image was composited from five separate photographic negatives (Robinson, 1860), the exact composition remains unknown and therefore it was not possible to accurately specify manipulated areas.

not. However, we note that the average accuracy rate of participants in this experiment (56.0%) compares favourably to the authors' results in a previous experiment (Caldwell et al, 2015) in which participants with only a brief familiarisation with principles of image manipulation achieved a lower mean accuracy rate of 37.5%.

Even when participants were told directly in the case of Images 8 and 13 that a feature had been added (missile and girl added), the ability to then identify which missile (missile number 3) and which girl (the girl in front) had been added was poor. For example, after participants of the cohort viewing the manipulated images were told an additional girl had been spliced into Image 13, their accuracy of what had been manipulated increased from 11 out of 40 participants to 28 out of 40 participants, but 11 participants nominated other girls in the image as spliced in as well.

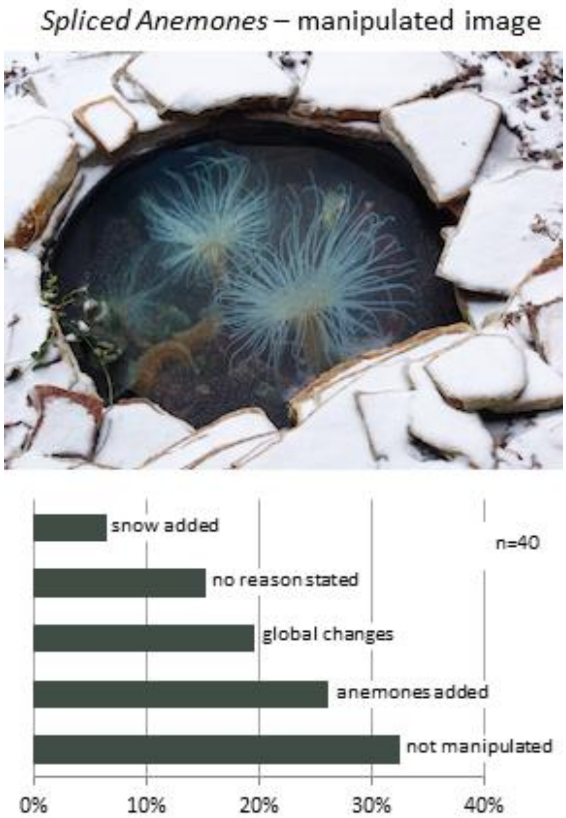


Figure 3-16: 67% of participants accurately identified this image as manipulated. 26% accurately identified that the anemones were added.

4. 1. Grounded Theory

We consider grounded theory outcomes at two levels: individual images and overall.

At an individual level, the characteristics of each image (semantics and content) yielded varying outcomes aligned with the nature of the image itself. Elements of interest to participants differed for each image, and corresponding determinations of what had been changed in images they stated had been manipulated, were also image-centric.

For example, in Figure 3-16, in which sea anemones were spliced into an image of a snow-covered pond, descriptions of manipulations focused on the anemones, the snow, and global changes such as sharpening and colour filtering. By contrast, in Image 11 in which an image of Queen Elizabeth II was spliced into a media scrum near Australian Prime Minister John Howard, manipulation descriptions focused on the Queen and John Howard. For participants identifying that an image was manipulated, the main descriptions of what they perceived as being changed are listed in Table 3-9 below.

Table 3-9 What participants stating image was manipulated perceived as being manipulated:

Image		n	Unmanipulated version	Manipulated version
Common to both cohorts	1	50	chameleon added/resized (23), background changed (19), match enlarged (4)	not presented
	2	21	colourised from black & white (10), global changes (7)	not presented
	3	45	not presented	made grayscale (32), other global changes (4)
	4	19	resized (5), cropped/zoomed (3)	not presented
	5	7	global changes (3)	not presented
	6	40	not presented	sepia/brown filter (22), man at window added (3)
	7	23	zebra added (11)	not presented
	8	40	not presented	missile(s) added (15), smoke added/changed (9)
	9	61	not presented	cow added (49), license plate blurred (9)
Differentiated	10	18	global changes (4)	global changes (6), cropped (2)
	11	29	global changes (3)	Queen added (10), global changes (3)
	12	47	global changes (13)	anemone added (12), global changes (9), snow added (3)
	13	30	global changes (6), red-eye reduction (5)	girl added (11), red-eye reduction (2)
	14	22	sign changed (6), woman added (4)	sign changed (13), woman added (4)

Some themes that arose in individual image descriptions were over-arching across all images presented. In particular, the themes of uncertainty and logic presented themselves in all images. These two aspects of participants’ responses are detailed in Table 3-10 below in which the number of instances of uncertainty and use of logic is shown. A response was flagged as using logic only if the logic process was evident in a participant's verbal response. For example, many participants stated that the chameleon in Image 1 was resized to a tiny size, by which it could be inferred that logically they believed that it was not possible for the chameleon to be that small, however only those who stated that logic aloud were counted. Similarly, while a long delay in responding might imply uncertainty, only those who stated they did not know or used other words of uncertainty in their responses were counted.

Table 3-10: Participants expressing overall themes of Uncertainty and Logic

Image	n	Uncertainty	Logic
1	80	24	14
2	80	12	5
3	80	13	5
4	80	14	3
5	80	10	3
6	80	11	10
7	80	5	3
8	80	20	12
9	80	22	16
10	40	14	3
11	40	16	9
12	40	13	10
13	40	9	5
14	40	12	11

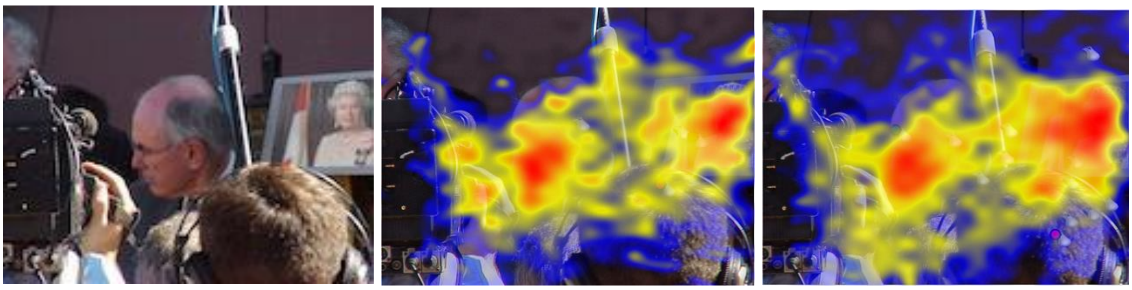
Participants often expressed their uncertainty by stating that they did not know if an image was manipulated, and making statements of general uneasiness with the image, evidenced by words such as 'perhaps, maybe' and phrases like 'I think,' 'I believe,' 'not sure,' 'doesn't look right.'. Uncertainty was also expressed in overt guessing.

Participants often employed logic (rightly or wrongly) in determining the validity of an image. For example, Image 4 was often (26%) noted as resized because it was pixelated rather than understood to be a low resolution photograph. In Image 13 (manipulated version) the girl was identified as added by 5 participants because she wasn't smiling like the others or because her lighting seemed different. In Image 1 the chameleon was stated as added because 'lizards aren't that small,' and in Image 7 the zebra was noted as added because 'zebras don't belong in a lingerie shop.'

The logic participants used in determining manipulations was informed by their pre-existing knowledge and beliefs. For example, for Image 9, a participant with a farming background commented "I know how much a cow weighs and a car would not stand up to that weight." In the case of Image 11, only 26 out of 80 participants seemed to know that the central figure being interviewed was former Prime Minister John Howard; with others referring to him as 'the man in the middle,' or 'the bald-headed man.' This may have been a function of the relative youth of the participants and the fact that John Howard has been out of office for over 7 years; the average age of those who identified the ex-Prime Minister was 27.7 (SD 11.9) in comparison to experiment average of 24.4 (SD 8.7) years. In the case where a colour image (Image 2) was filtered to create a black & white version (Image 3) participants with photo editing knowledge correctly identified it as manipulated with the use of appropriate language (desaturated, filtered) to describe the change.

However, the use of logic was counterproductive if the logic was not sound. In the case of Image 1, for example, 5 participants stated the image was manipulated based on their belief that match heads cannot be black in colour, which is not the case.

4. 2. Eye Gaze



A **B** **C**

Figure 3-17: Comparing 'heatmap' of eye gaze intensity of participants presented with Image 11
A) Image 11 as presented, **B**) Eye gaze of participants stating: unmanipulated, **C**) Eye gaze of participants stating: manipulated. *Key: Red - most viewed, yellow - less viewed, blue - least viewed.*

In the case of Images 10, 11, and 12, participants' eye gaze indicated that increased viewing time as denoted by a greater number of fixations and durations of eye gaze were associated with a greater rate of accuracy in identifying manipulating issues. In addition, although the remaining 5 images did not show this outcome, an overall effect of increased attention associated with increased accuracy was noted when all images were taken as a whole. It is not clear whether this was the result of a person suspecting a manipulated region at the commencement of viewing an image and extending their attention as a result to seek corroboration with eye gaze or suspecting the manipulation as a result of extended attention of the target area.

4. 3. Grounded Theory and Eye Gaze

In attempting to understand the reasons that Images 10, 11, and 12 demonstrated a significant difference in attention via eye gaze fixations and durations, we considered the nature of the images. Some common motivations for manipulated images are personal, humour, politics and fraud. It may be that in comparison to more 'mainstream' images that showed a greater similarity in attention (Images 3, 8, 9, 13, 14), Images 10-12 may have been 'harder work' to interpret. For example, although Image 11 may be a political image, it contains an unusual juxtaposition of John Howard and an image of Queen Elizabeth II (Figure 3) that seemed to confound participants, especially those who stated the image was manipulated.

These participants devoted more of their attention and verbal responses to the image of the Queen than to John Howard. This effect on visual attention is illustrated in the heat map of gaze intensity in C of Figure 3-17, which is greater for the image of the Queen than for John Howard, and also greater in comparison to the same location in image B. The 'unmanipulated' group represented in B of Figure 3 also referred to the image of the Queen 23% less in their comments than the 'manipulated' group shown in C.

It is interesting to note that Image 6 was almost universally misunderstood. Of the 80 participants who viewed this early version of image manipulation from the 19th century, only 3 came close to correctly interpreting that the entire image was a composite derived from 5 negatives; they noted that the man at the window had been added, although this represents only a part of the composition. This may indicate a 'hiding effect' in which a more significant perceived, though incorrect, manipulation (filtering) 'hides' the compositing. This effect is more overt in the case of Image 9: although 49 participants out of 80 stated that the cow has been added to the bonnet of the BMW, only 9 noted in addition that the license plate had been obscured as well, and indeed the blurred license plate received only 2.9% of eye gaze attention by all participants.

5. Conclusion

Based on the outcome of this experiment, we conclude that being presented with an image is in itself insufficient to reliably identify if the image we view has been altered. With an average success rate of 56.0%, we get it wrong 44% of the time. Even when we do successfully identify manipulated images, it is far more difficult for us to tell *what* has been manipulated (27.7%). However, eye gaze tracking shows that for some images and as an overall trend, increased attention (fixation, duration of eye gaze) to manipulated areas of photographs tends to be associated with greater accuracy in deciding if an image is manipulated.

Our eye gaze is a partial reflection of the features of an image that we non-consciously note, but this is not always a predictor of conscious accuracy. Further, more prominent features of an image may obscure our recognition of less obvious manipulations.

In the absence of additional knowledge about the photograph, we resort to the use of logic derived from personal experience which may or may not be relevant, and remain uncertain of our conclusions.

This is problematic because we use images in almost every walk of life from social media to advertising, and from news images to health information, without being able to determine if and how these images have been manipulated.

If viewing a standalone image is not sufficient to allow us to identify if it has been manipulated, even after being pre-familiarised with concepts of image manipulation, then how can we interpret images correctly? It seems plausible that it is necessary for images to be accompanied by a source of additional information such as one or more of an assertion of the status of the image (manipulated or not and how), metadata, context, reference images, or verbal description. We will be investigating this in our further work. If users can be truthfully convinced of the veracity of the images they view online, then this has implications for the design for many forms of human computer interaction via the web.

- end of paper -

3.3 Image authentication framework

A cornerstone of the research reported in this thesis is the image authentication framework I have developed. The design is informed by my research. It is a cross-disciplinary amalgam of the exigencies of human perception of manipulated and unmanipulated images, and the technologies that can deliver images to humans in a package of information and functionality that allows them to understand the meaning and relationship to real people, places and events.

The initially posited self-contained mobile image format remains core to the framework, however as perceptions of and perspectives on manipulated images came to light in the research, the .msci format became enmeshed in broader contexts of communication, artistic interpretation, trust and security. The figure below presents a high level view of the framework developed as an outgrowth of understanding of the way people interact with manipulated images.

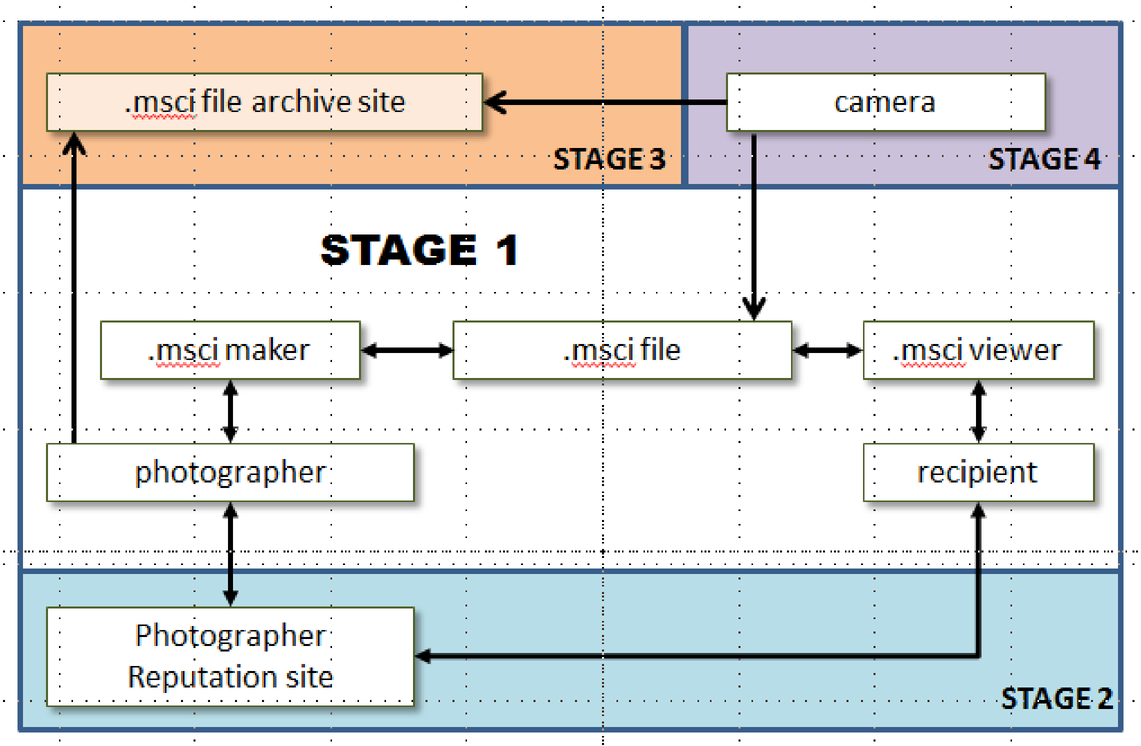


Figure 3-18: Top level image authentication framework

The solution proposed to address the issue of image authentication is a technical and procedural framework that fuses technology and social solutions. It recognises that image manipulation is now firmly enmeshed in the discipline of photography, and has many important benefits, and therefore any photographic authentication framework must (counterintuitively) accommodate manipulated images, while still firmly focusing on unmanipulated photographs.

At the heart of the solution is the .msci file format which presents images and associated information in a standalone format. Technical aspects that are new in this format are:

- the idea of ‘packaging’ original (proof) and modified (presentation) versions of the image;
- extraction of metadata from the obscure locations in which it currently resides to a clearly visible display; and
- a difference map to highlight changes.

The framework as a whole also does something novel that is not technical but social - it harnesses the power of social mores, in that it creates a positive assertion on the part of the distributor that the original image is original and not modified. This aspect engages the photographer with issues of honesty, reputation, artistry, meaning production, history, and many other elements of our social fabric. While there will always be individuals who break the rules, the vast majority of producers of digital photographs adhere to these social conventions, and this file format offers them a way to demonstrate how their work represents, reflects, or adds to our understanding of the world. For consumers of digital images, this format elevates both the meaning inherent in the image(s), and their ability to have confidence in their relationship to the real world.

Further explication of the framework is provided in 5.7, and detailed in *Appendix B: Design Specifications for Image Authentication Framework including Content and Semantics for .msci file format*.

4 Neural net eye gaze pattern classification

A question arises when comparing experiment participant responses to the question "Is this image manipulated or not manipulated?" to the data recorded as their eyes viewed the image. Is it possible to establish a computational algorithm that can distinguish between different eye gaze patterns as they relate to manipulated and unmanipulated images? The data does not lend itself to straightforward processing; rather, such an algorithm would need to 'find its way' through the dataset to discern commonalities for features of the eye gaze data that fit the data samples to one or another outcome. To attempt to distinguish if eye gaze patterns can be differentiated between different outcomes, I employed pattern matching artificial neural net (NN) analysis using the feed-forward error back-propagation algorithm. This appears to be a new application of neural net analysis as I was unable to locate anything relevant in the literature in respect of applying neural networks to tampered images.

4.1 Differentiated image cohorts from Experiment B

A natural starting point for this analysis was Experiment B, in which subjects were divided into different cohorts and shown 5 images that were manipulated for one cohort and unmanipulated for the other. The purpose for this part of Experiment B was to compare eye gaze and verbal responses to the two types of images, and therefore the eye gaze data were already divided into the two desired groups.

Eye gaze data were prepared by summarizing number and duration of fixations inside and outside the target manipulated regions, and these data formed the feature set for each participant for each image. The output data comprised a binary feature set in which 1 denoted that that feature belonged to a manipulated image, and 0 denoted that the feature belonged to an unmanipulated image. The NN was then used to attempt classification of eye gaze patterns against their status as manipulated or unmanipulated for each of the 5 differentiated images (Figure 4-1) separately as well as together. Thirty passes for each dataset was executed and the averages charted (Figure 4-2).

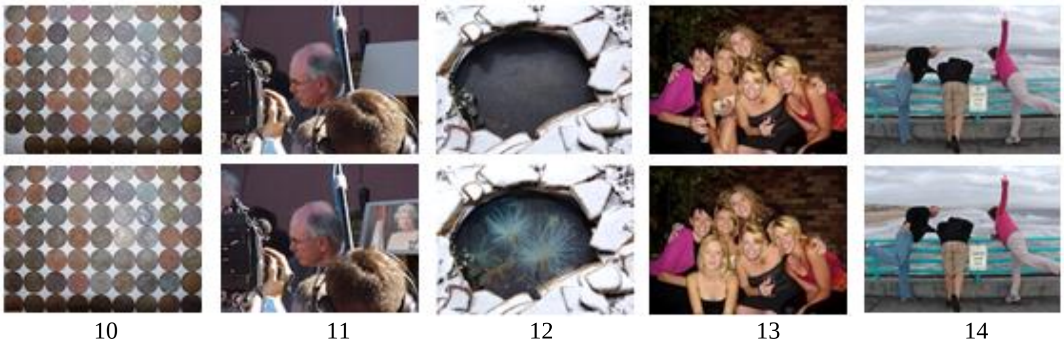


Figure 4-1: Differentiated images from Experiment B, unmanipulated (top row) and manipulated (bottom row)

The NN was able to distinguish between the eye gaze patterns of participants viewing the manipulated versions of the images and those viewing the unmanipulated versions (Figure 4-2). Significance was calculated using a 1-tailed T Test comparing the results for each image against chance (0.5). Significance is p<0.01 in all cases.

Accuracy varied in these results, with the NN better able to distinguish patterns between the eye gaze of participants for images 11, 12, and 13, and less well for images 10 and 14.

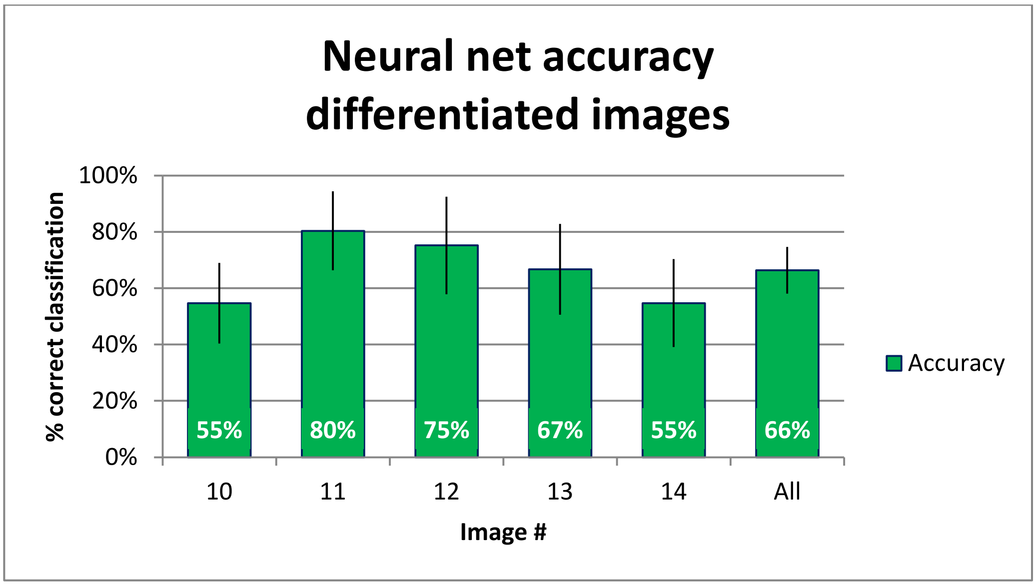


Figure 4-2 Neural net accuracy in classifying eye gaze patterns of subjects viewing differentiated manipulated and unmanipulated images from Experiment B. n(5)=76, n(9)=75, n(10)=76, n(12)=74, n(15)=75 α for results = p<0.01

These accuracy levels were reflected in the number of correct outcomes vs incorrect outcomes for classifications NN 'votes' for images. For example, the average number of correct outcomes for image 11 (80% accuracy) was 61 out of 75, while for image 10 (52% accuracy) there were 39 out of 76 average correct outcomes (Figure 4-3).

Table 4-1: NN votes

	10	11	12	13	14	All
true negative	23	32	30	27	20	27
false negative	21	9	11	14	17	15
negative subtotal	***44***	***41***	***41***	***41***	***38***	***41***
true positive	16	29	27	22	21	23
false positive	16	5	8	11	17	11
positive subtotal	***32***	***34***	***35***	***33***	***37***	***34***
totals	***76***	***75***	***76***	***74***	***75***	***75***

Outcomes can be further teased out into 4 categories: true positives, false positives, true negatives and false negatives (Table 4-1 and Figure 4-3). From this view, it appears to be more likely for the NN to vote for eye gaze behavior as relating to the unmanipulated state (average 41 out of ~75) rather than the manipulated state (average 34 out of ~75) even though both states were equivalently represented in the experiment results even after elimination of non-qualifying views.

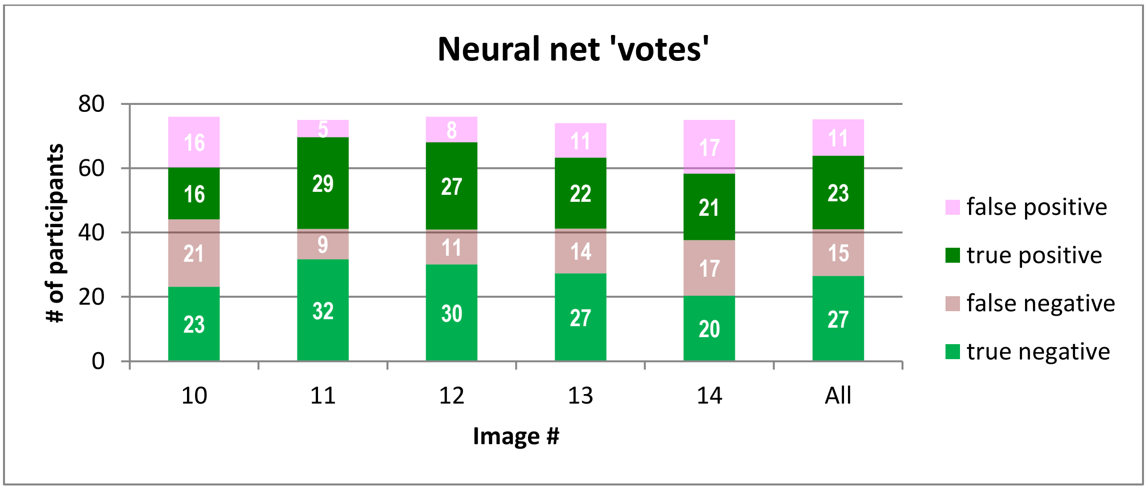


Figure 4-3: NN assignment of participants' eye gaze to image manipulation state

The analysis was also run with all sample data and targets for the differentiated images with an accuracy result of 61.9% (SD 8.3%, α = p<0.01). This is different from the simple average of the results of each of the 5 differentiated analyses noted in Figure 4-2, which is 65.7% (SD 15.4%). This may indicate that classification success is more difficult when the feature set spans the range of all 5 images.

A question that then arises from these data is how the accuracy of the neural nets compared to experiment participants' nominations of manipulated versus manipulated. In other words, how did the NN votes compare to the participants' votes?

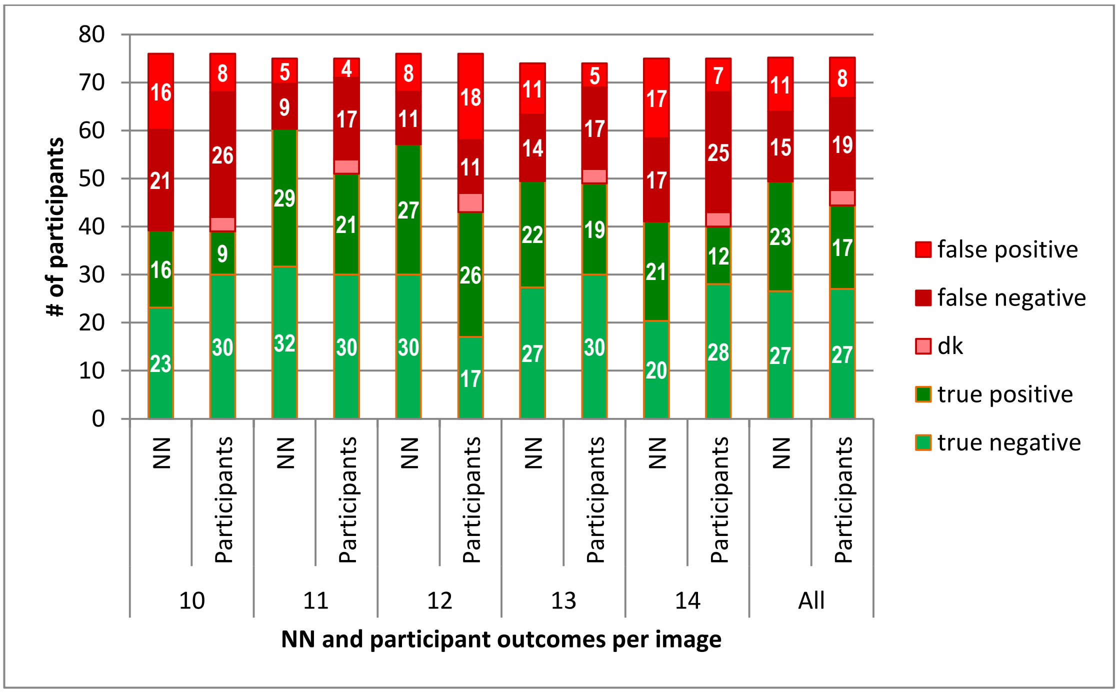


Figure 4-4: Performance of pattern-matching NN vs participant choice

The pattern matching NN compared favourably to participants' choices (Figure 4-4 and Table 4-2). In three of the five cases (images 10, 13 and 14), the NN had a combined accuracy of true positive and true negative votes commensurate with participant votes. In two cases, images 11 and 12, the neural net votes were more accurate than participant votes. On average, the NN performed better (49 out of 75, or 65.3%) than participants (44 out of 75, or 58.7%).

Table 4-2: Comparison of NN and participant votes

	10		11		12		13		14		All	
	NN	Partici-pants	NN	Partici-pants	NN	Partici-pants	NN	Partici-pants	NN	Partici-pants	NN	Partici-pants
true negative	23	30	32	30	30	17	27	30	20	28	27	27
true positive	16	9	29	21	27	26	22	19	21	12	23	17
accuracy subtotal	*39*	*39*	*60*	*51*	*57*	*43*	*49*	*49*	*41*	*40*	*49*	*44*
dk	0	3	0	3	0	4	0	3	0	3	0	3
false negative	21	26	9	17	11	11	14	17	17	25	15	19
false positive	16	8	5	4	8	18	11	5	17	7	11	8
inaccuracy subtotal	*37*	*37*	*15*	*24*	*19*	*33*	*25*	*25*	*34*	*35*	*26*	*31*
Total	76		75		76		74		75		75	

There were, however, differences in how the NN and the participants identified whether the image was manipulated or not. The NN was not allowed to choose 'don't know' as some participants did despite prompting to choose one or the other.

With the exception of image 12, the participants viewing the unmanipulated images performed better than the participants viewing the manipulated images. The NN votes were more evenly spread across the two types of images. This difference could be seen in the false positives and negatives as well, most obviously in image 14, where the NN voted for 17 false negatives and 17 false positives while participants voted for 23 false negatives and only 7 false positives. Discrepancies such as these may reflect a human bias to believe in images even after being trained in image manipulation.

4.2 Undifferentiated image cohorts from Experiment B

Participants in this experiment also viewed 9 images that were manipulated but not differentiated between cohorts (Figure 4-5). This was a useful data set to investigate whether people look at images differently in determining their veracity.

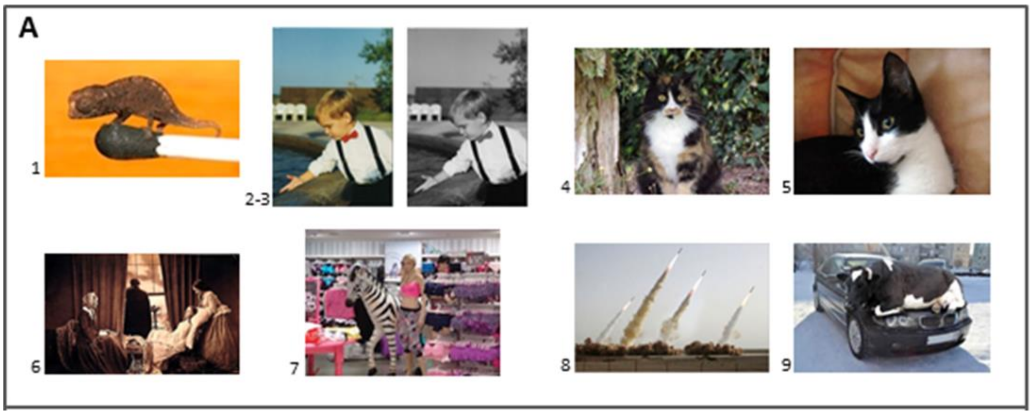


Figure 4-5: Experiment B images viewed by both cohorts
Unmanipulated images are 1, 2, 4, 5, 7; manipulated images are 3, 6, 8, 9

The pattern net algorithm was again used, but this time the classification was not to determine whether the eye gaze pattern was attributable to a manipulated or unmanipulated image, but rather to match participant eye gaze to their decisions. In each case the status of manipulation is known for each image and is identical for all participants. Therefore the feature set is the same as in the previous analysis, however the outputs are participant *decisions* as to whether the image has been manipulated or not rather than the actual status of the image. This means that the NN is attempting to learn how to distinguish between the eye gaze patterns of the participants who thought an image was manipulated vs the eye gaze of the participants who thought it was not.

For each image the cohort is nominally 80 subjects, though as before some views had to be excluded due to eye gaze tracking issues. Further, because any ‘don’t know’ responses were not consistent with valid targets for the NN, the data for participants with ‘don’t know’ responses were excluded.

Again, the NN again was able to distinguish between the eye gazes of participants who chose differently (Figure 4-6).

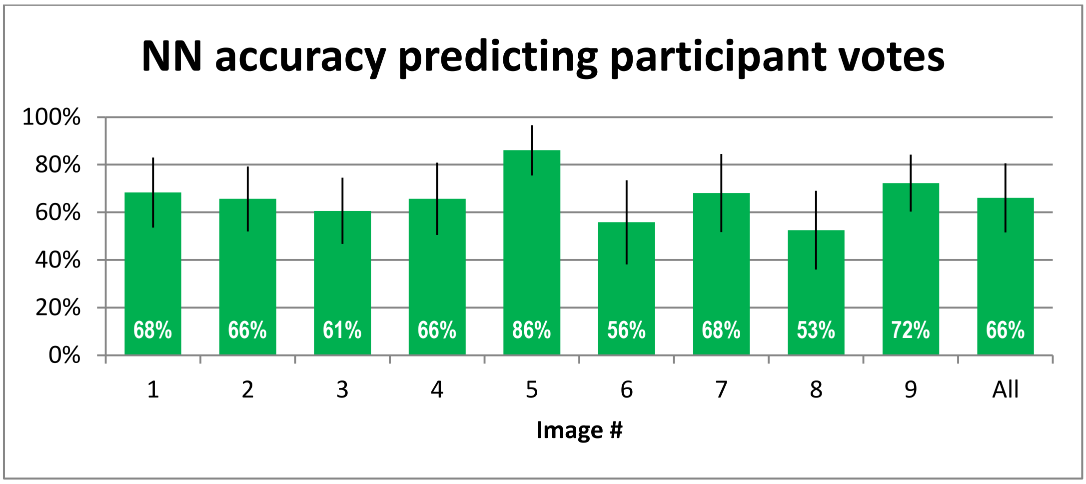


Figure 4-6: NN accuracy predicting participant votes for undifferentiated images
n(1)=65, n(2)=71, n(3)=65, n(4)=60, n(5)=60, n(6)=75, n(7)=72, n(8)=69, n(9)=73
α for all results = $p<0.01$

Overall, the results showed a 66% accuracy in correctly identifying the participants' votes through NN analysis of their eye gaze tracks (SD 0.15). Since every participant viewed the same images, this implies a factor or factors present and identifiable to differing degrees in their eye gaze tracks that corresponded to their ultimate choice.

As with the differentiated images, outcomes could be further teased out into 4 categories: true positives, false positives, true negatives and false negatives (Table 4-1 and Figure 4-3). From this view, it appears to be more likely for the NN to vote for eye gaze behavior as relating to the unmanipulated state (average 41 out of ~75) rather than the manipulated state (average 34 out of ~75).

Table 4-3: NN classifications of participant eye gaze

	1 (u)	2 (u)	3 (m)	4 (u)	5 (u)	6 (m)	7 (u)	8 (m)	9 (m)	Av All
Participant votes – neg	18	37	21	54	37	37	49	29	17	33
Participant votes - pos	47	38	44	7	38	38	23	40	56	37
true negative	4	19	3	44	19	19	42	9	1	18
false negative	7	16	7	6	16	16	16	13	4	11
subtotal negative	12	35	10	50	35	35	58	23	5	29
true positive	40	22	37	1	22	22	7	27	52	25
false positive	14	18	18	9	18	18	7	20	16	15
subtotal positive	53	40	55	10	40	40	14	46	68	41
totals	65	75	65	60	75	75	72	69	73	70

From Table 4-3 we can see that image 1 (the chameleon on the matchstick) ‘fooled’ the most people, with 47 of 65 saying that it is manipulated when it is not. Participants did best with image 9, the image of the Holstein cow laying upon the hood of a BMW, where 56 of 73 said it was manipulated and it is.

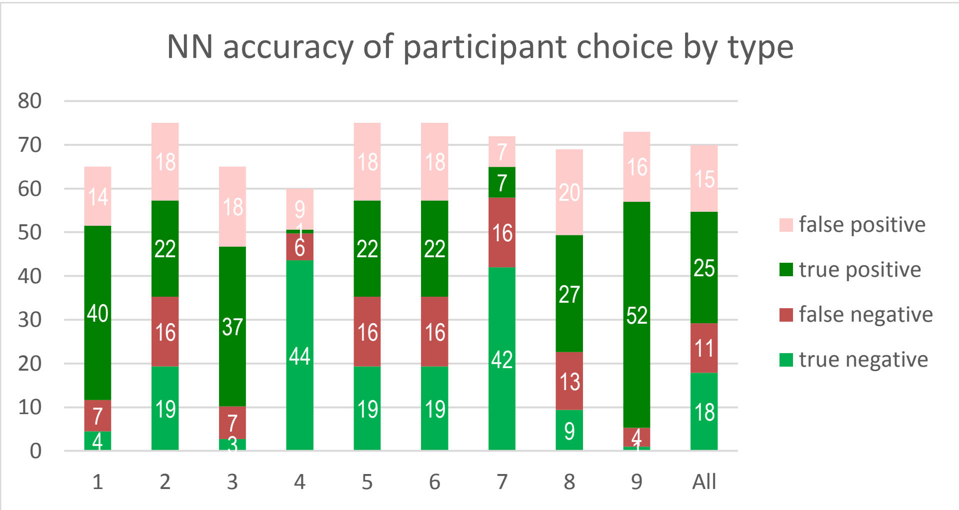


Figure 4-7: NN accuracy of participant choice (vote) by type for undifferentiated images

This information is a little difficult to interpret. However, remembering that a) the unmanipulated images are 1, 2, 4, 5, 7 and the manipulated images are 3, 6, 8, 9, and b) the NN is attempting to classify participants eye gaze against the participants’ determinations of whether the image was manipulated or not (not whether the image itself was manipulated or not), we can work it out. What this chart tells us is that in almost all cases it was possible for the NN to see differences between the participants’ eye gaze aligning with their vote.

In particular, the NN performed best on images 1, 3, 4, 7, and 9, while the others were more equivocal. For example, the result for image 1, the chameleon on the matchstick (unmanipulated), indicates that the NN strongly identified the eye gaze patterns of those who would say it was manipulated (40 of 65). In contrast, image 2, the colour photograph of a boy at a fountain (unmanipulated), results indicate that it was more difficult to classify the eye gaze into the two groups.

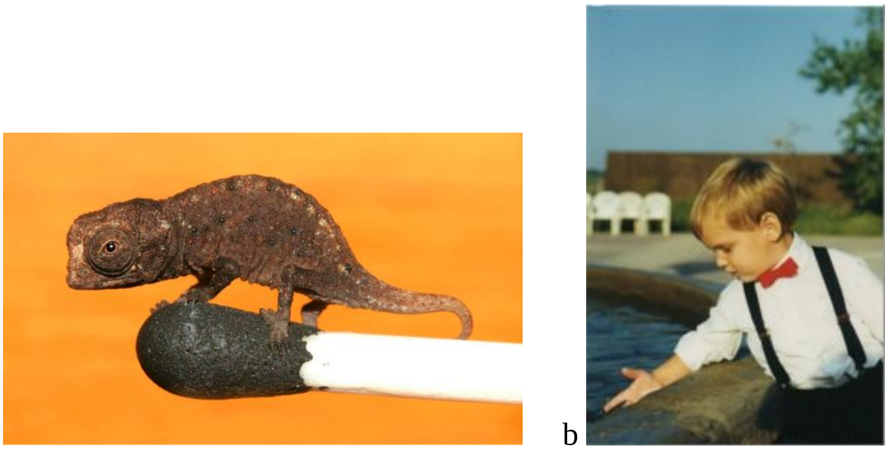
a  b

Figure 4-8: Best performance by NN for unmanipulated image (a) and worst (b)

If we revisit the idea of whether the image was manipulated or not, other signals emerge. Of the manipulated images, it appears that the NN was more easily able to identify those who would say that the image was manipulated (true positive) for image 9, and least able for image 6. Of the unmanipulated images, the best performance (true negative) was for image 4 (Figure 4-9), with images 2 and 7 tying for worst NN performance.

a  b

Figure 4-9: Best performance by NN for manipulated image 9, (a) and unmanipulated image 4, (b)

4.3 Zone-based NN analysis of eye gaze as a predictor of participant votes

Having considered how effective pattern matching neural net analysis was in identifying eye gaze patterns as classifiers of participant votes, it seemed useful to take this a step further and attempt to refine how image elements informed these classifications. What are the important parts of an image (as judged by our eye gaze as we look at it) that cause us to choose if an image is manipulated or not, and that the NN uses in classifying our votes? And which features of eye gaze (frequency or duration) are more important?

To investigate this, I separated the eye gaze data into three different feature sets:
Set 1 = count of fixations and sum of durations
Set 2 = count of fixations, count of samples in durations and sum of durations
Set 3 = count of samples in durations and sum of durations

These feature sets were applied to a range of zones for three of the images from experiment B:
Image 1 – Chameleon on a matchstick {screen, image, chameleon, matchstick}
Image 7 – Zebra in a lingerie shop {screen, image, zebra, mannequins}
Image 8 – North Korean missiles {image, missile 1, missile 2, missile 3, missile 4}

The count of fixations relates to the frequency with which a participants' eye gaze fixates in the zone in question. The count of samples results from Eyeworks counting eye gaze durations in a fixation on a timed basis, which in this case was 1.5 milliseconds. The sum of durations is closely related to counts of samples but is measured as a total quantity of time spent viewing the zone.

4.3.1 Zebra in a lingerie shop

All participants viewed image 7, an unmanipulated photograph of a lingerie shop within which the owners had placed a large fiberglass zebra presumably for attracting customers into the shop. I intuitively surmised that participants who paid more attention to the zebra would be more likely to say the image was unmanipulated, perhaps by recognizing it as a sculpture rather than a real animal.

Figure 4-10: Zebra in a lingerie shop

Further, the fact that the women standing near the zebra were also not real (mannequins) might have played a part in participants' voting. To test if the attention paid to the zebra, the mannequins, or both yielded different success rates, the NN for undifferentiated images used in 4.2 was run using a range of combinations of the relevant feature sets (fixations and durations of eye gaze used for each zone noted). The data sets and results are listed in Table 4-4 (all $\alpha < 0.01$).

Table 4-4: Image 7 feature set and results for zones of interest

	Accuracy (%) feature sets 1-3						**Zones of interest**			
Set	**1**	**(sd)**	**2**	**(sd)**	**3**	**(sd)**	**Screen**	**Image**	**Zebra**	**Mannequins**
O	**68.1**	16.4	66.5	15.4	67.8	14.7	✓	✓		
A	66.6	14.6	66.8	15.5	**67.8**	14.7	✓	✓	✓	
B	*68.4*	14.3	68.0	15.2	**68.7**	15.5	✓	✓		✓
C	66.0	16.7	67.7	16.6	**68.2**	15.1	✓	✓	✓	✓
D	67.0	15.9	*68.3*	14.6	***70.3***	12.3		✓	✓	
E	68.0	15.9	68.5	15.4	**68.9**	14.4		✓		✓
F	67.4	15.8	67.4	16.5	**68.2**	15.2		✓	✓	✓
G	63.1	15.2	64.1	15.9	**65.5**	14.6			✓	✓
H	65.9	14.8	65.7	14.8	**66.0**	14.8	✓		✓	
I	**66.0**	14.3	65.8	15.4	65.5	14.5	✓			✓

(Greatest accuracy feature set for each set O,A-I **bolded** and greatest accuracy for each feature set *italicized*.)
Feature sets contain:
Set 1 = count of fixations and sum of durations
Set 2 = count of fixations, count of samples in durations and sum of durations
Set 3 = count of samples in durations and sum of durations

The first thing that is noticeable is that the NN was able to predict participant votes from at least 63.1% to up to 70.3% accuracy. This is in line with analyses in 4.2. This means that for almost any case, we can guess with between a 6-7 out of 10 likelihood of success whether a person thinks an image is manipulated or not, just by looking at their eye gaze data.

The second thing that is noticeable is that the feature set 3 – containing the sum of samples in durations and the total sum of durations most accurately predicted the participants vote in almost all of the zone sets interrogated (zone set O, or the original set tested, and one set I were the exceptions). Of the three sets, this had the strongest relationship to the length of time spent looking at the zones, and the weakest relationship to the frequency with which the zones were viewed. This appears to indicate that in the case of this image, the NN was best able to predict participants' votes based on the quantity of attention *time* rather than the quantity of attention *events*.

There was little difference between the effectiveness of the zebra and mannequins zones in determining participant votes, with an average 66.78% accuracy for the former, and 66.97% for the latter.

4.3.2 Chameleon on a matchstick

The chameleon photograph (Figure 4-8) confused most people. Of the 12 participants of experiment A and 80 participants of experiment B, 62 believed the photograph was a fake, when it was not. Because this image is relatively simple, with just a small animal, a matchstick and a background, it was a good candidate to see if it is possible to identify whether one visual element was more important than another in participant voting.

Table 4-5: Image 1 feature set and results for zones of interest

	Accuracy (%) feature sets 1-3						Zones of interest			
Set	**1**	**(sd)**	**2**	**(sd)**	**3**	**(sd)**	**Screen**	**Image**	**Chameleon**	**Matchstick**
O	68.3	14.7	69.3	13.9	**69.4**	14.3	✓	✓		
A	69.0	13.8	68.2	16.3	69.0	14.2	✓	✓	✓	
B	68.0	14.6	**68.5**	14.8	68.2	14.0	✓	✓		✓
C	66.9	16.5	67.3	15.6	**68.9**	14.8	✓	✓	✓	✓
D	69.5	14.5	70.0	15.4	***70.3***	15.2		✓	✓	
E	69.5	15.7	***71.2***	14.3	69.7	14.8		✓		✓
F	***69.8***	15.3	68.7	16.2	69.5	14.4		✓	✓	✓
G	66.9	15.2	67.7	13.9	67.7	14.8			✓	✓
H	66.0	15.1	**67.8**	13.2	67.6	13.3	✓		✓	
I	66.5	15.1	**68.9**	12.1	66.1	14.2	✓			✓

(Greatest accuracy feature set for each set O,A-I **bolded** and greatest accuracy for each feature set *italicized*.)
Feature sets contain:
Set 1 = count of fixations and sum of durations
Set 2 = count of fixations, count of samples in durations and sum of durations
Set 3 = count of samples in durations and sum of durations

Again, the NN was able to discern between the eye gaze patters of participants who decided the photo was manipulated versus those who did not – in this case to a 66.0% to 71.2% level. Unlike the previous example in 0, there was not a stronger signal in feature set 3, implying that duration of view was less important in deciding outcomes in this case.

Participant gaze on the chameleon or the matchstick did not affect NN accuracy; the average accuracy for the NN for all feature sets using the chameleon zone was 68.38%, and for the matchstick zone it was 68.33%.

It seems that for both of the above images, participants' eye gaze in specific zones that might be considered manipulations (zebra, mannequins, chameleon, matchstick) have relatively equal weights as far as the NN is concerned in classifying participant votes. This prompts the question, can eye gaze tell us *what* has been manipulated in an image?

4.4 Zone based NN analysis of eye gaze as a predictor of manipulations

To determine if participants' eye gaze data can provide clues to the areas of manipulation, I selected image 8, the image from North Korea of a missile launch. In addition to having discrete objects in the image, participants in the experiment were specifically told that one of the images had been cloned into the photograph and asked which one they thought it was. This provided extra outputs against which the NN could attempt to classify the eye gaze data.

4.4.1 Missing missile from North Korean launch

It is known that in the photograph of 4 missiles launched in a test by North Korea, only three missiles launched successfully, with one of the missiles having been cloned in using photo editing (The Telegraph UK, 2016). A short perusal of the image shows that the missile appearing in the third from left position has the same dust cloud as the fourth missile, and the same vapour plume as the second missile. However, in experiment A, no participant identified the image as manipulated, and in experiment B, where the image was used again this time with training in image manipulation, 38.9% of participants were unable to identify it as manipulated, and only 11 participants (out of 70 valid responses) were able to identify the cloned missile correctly, even after being told that one of the missiles had been cloned into the photograph and then being asked

Figure 4-11: North Korean missile launch with faked missile (3rd from left)

which one they thought it was. Despite having a 1 out of 4 chance (25%) at getting it

right by sheer luck, the success rate was only 15.7%.

In the first instance, I ran the NN analysis in respect of participant votes as to whether or not the image was manipulated. Fixation counts and durations for each of the zone combinations in Table 4-6 were used as samples in the same combinations as the above examples at 4.3.1 and 4.3.2. All possible combinations of zones delimiting the missiles within the image were tested.

Table 4-6: Image 8 feature set and results for zones of interest

	Accuracy (%) feature sets 1-3							Missile zones of interest			
Set	**1**	**(sd)**	**2**	**(sd)**	**3**	**(sd)**	**Image**	**M 1**	**M 2**	**M3**	**M4**
A	53.7	15.8	*56.0*	16.2	***56.5***	15.4	✓	✓			
B	54.7	17.1	**55.0**	16.2	53.6	14.7	✓		✓		
C	**54.4**	16.2	53.9	15.3	53.0	15.4	✓			✓	
D	**54.6**	16.6	54.3	14.9	54.1	16.1	✓				✓
E	**54.7**	15.8	52.4	15.9	52.4	15.9	✓	✓	✓		
F	52.9	16.1	**53.2**	17.0	50.6	16.0	✓	✓	✓	✓	
G	**54.6**	14.8	54.3	15.8	53.2	16.4	✓	✓	✓	✓	✓
H	**54.4**	15.2	52.8	16.2	52.9	14.7	✓		✓	✓	
I	52.8	16.6	**53.1**	16.9	52.2	15.8	✓		✓	✓	✓
J	***55.1***	16.8	53.8	15.4	53.3	16.0	✓			✓	✓
K	52.5	16.0	**54.3**	15.4	53.7	16.1	✓	✓			✓
L	**54.8**	17.7	54.6	15.4	51.5	16.0	✓	✓		✓	
M	53.5	15.2	53.0	15.7	**55.1**	15.0	✓		✓		✓
N	53.1	15.6	52.1	15.7	**53.6**	16.6	✓	✓		✓	✓
O	52.2	15.9	**53.6**	15.8	53.2	15.6	✓	✓	✓		✓

(Greatest accuracy feature set for each set A-O **bolded** and greatest accuracy for each feature set *italicized*.)
Feature sets contain:
Set 1 = count of fixations and sum of durations
Set 2 = count of fixations, count of samples in durations and sum of durations
Set 3 = count of samples in durations and sum of durations

These results indicate that Set 1, the count of fixations and sum of durations feature set is the most useful features grouping in identifying participant votes, with 7 of the highest accuracy outcomes out of 15. Feature set 2 was most accurate in 5 cases, and feature set 3 was most accurate in only 3 cases.

To identify whether participants' eye gaze data could be used to predict which missile the participant would pick as the manipulated area of the photograph, the pattern-matching NN was used to compare eye gaze data against each of the votes for each missile. Because feature set 1, number of fixations and sum of durations, was the most effective set based on the results set out in Table 4-6, this feature set was used for all tests in this step.

Table 4-7: Eye gaze comparison selecting missiles 1 - 4

	Accuracy (%) by participant vote for 'which missile'									Missile zones			
Set	M1	(sd)	M2	(sd)	M3	(sd)	M4	(sd)	Image	M 1	M 2	M3	M4
A	57.9	13.9	83.1	9.5	72.5	12.8	73.7	12.2	✓	✓			
B	56.5	14.3	84.0	9.2	71.6	13.5	74.4	12.6	✓		✓		
C	58.8	12.9	83.9	9.5	72.6	12.1	72.6	13.4	✓			✓	
D	58.1	16.8	**84.5**	8.5	71.8	13.8	**74.6**	11.8	✓				✓
E	55.7	15.5	82.7	10.1	70.1	14.4	73.9	12.6	✓	✓	✓		
F	58.5	13.6	84.0	9.1	71.4	14.0	73.2	12.0	✓	✓	✓	✓	
G	58.0	15.0	82.7	10.2	71.0	14.3	71.9	14.6	✓	✓	✓	✓	✓
H	57.2	14.7	84.1	8.8	**72.9**	12.1	73.8	13.3	✓		✓	✓	
I	59.0	14.7	84.3	8.8	71.3	13.5	73.3	13.7	✓		✓	✓	✓
J	58.8	15.1	83.8	8.9	71.1	12.8	74.3	12.0	✓			✓	✓
K	58.7	15.1	84.1	8.4	72.2	12.7	73.3	12.7	✓	✓			✓
L	58.4	13.9	84.2	8.5	71.2	13.2	74.2	11.6	✓	✓		✓	
M	**59.2**	15.0	**84.5**	8.1	72.3	11.6	74.3	13.0	✓		✓		✓
N	57.8	15.0	83.7	9.8	71.7	13.4	73.5	12.5	✓	✓		✓	✓
O	58.6	13.2	83.7	9.1	72.1	12.7	73.6	13.0	✓	✓	✓		✓

N = 68 valid participants (Greatest accuracy feature set for each missile **bolded**)
Feature set used for all is Set 1 = count of fixations and sum of durations

It appears that it is more difficult for the NN to predict whether the experiment participant will identify the image as manipulated or unmanipulated than to classify *which* missile the participant will pick as added. The best result for the NN predicting whether the participant vote was for manipulated or unmanipulated is 56.5%, while the best results for predicting if the participant would say the missile that had been added is missile 1, 2, 3, or 4 are 59.2%, 84.5%, 72.9%, and 74.6% respectively (Table 4-7).

There is one more question worth putting to these data before we leave this image. Is it possible that the NN can discern the right answer in the eye gaze patterns of the experiment participants even when the participants choose incorrectly? In other words, do features of eye gaze manifest the thinking of participants at a level below normal conscious thought?

To test this, I ran the NN using 4 sample sets (1 each for each of the 4 missiles) per valid participant using the correct answer as targets for each set. The overall average success of the NN in matching participants' eye gaze to the correct answers was 75.1% (sd 4.2%). Earlier (4.4.1) we saw that the success rate for participants in choosing missile 3 as the missile that had been cloned into the image is 15.7%. This is considerably less accurate than the NN predicted from their eye gaze.

There are questions that are raised by all of the results in this chapter. What is the difference between the two categories of eye gaze relating to participant choice? Why are the predictive abilities of the NN so different for different images? Why is there a bias one way or another towards participant choice in different images? Why and how do we moderate the findings of our brain functioning in expressing conscious choices? Learning more about the answers to these questions will require additional NN analysis with varying feature sets, and further experiments. These activities would be well worth undertaking as future work.

However, it is clear that the information we receive visually is complex both in the actual data and the way in which we process it to arrive at conscious impressions of the images we see. We have earlier seen that additional visual cues in the form of comparison images and difference maps increase our success in detecting image tampering from 37% (unaided) to as high as 97% (with difference maps). We now see that neural networks can predict our decisions based on our eyegaze, which further demonstrates the importance of additional visual cues in our cognitive processes, and any image credibility solution framework needs to take this into consideration.

5 Discussion

This thesis has investigated a range of aspects of image credibility from the perspectives of human attitudes. I have surveyed the heterogeneous array of attitudes towards image tampering, our competence to recognize image tampering, and the rationality we apply in mentally processing the images as credible or not. Using eye gaze tracking, I investigated how we use our vision in ‘reading’ an image and identifying anomalies. Using artificial neural network analysis, I researched aspects of the relationship between how we use our vision to interrogate an image and our conscious decisions.

Overall, the research suggests that, on the whole, we do not as yet fully attend to the implications of image tampering, although awareness raising is effective.. When explicitly urged to consider whether an image is manipulated, and what has been manipulated, our ability to do it is respectively weak and very weak. Our rational minds both help and hinder us in critically thinking about image veracity, and some of us do not believe photographs are ever ‘real,’ preferring instead to assimilate the meaning in the image as an individual, photoartistic expression. We are more confident than we should be about our conclusions on the veracity of the images we view. When given tools to compare original and tampered images, we improve significantly, and when given with overt illustration of image differences we reach a very high level of understanding of the credibility of a given image. However, technological solutions are vulnerable to ‘hacking’ as seen in the failed efforts of camera manufacturers to implement technological solutions.

These findings inform and support the need for an image credibility framework such as the one proposed. They indicate that while visual literacy training can help us view images more critically, it does not solve the problem, but giving people comparison images and difference maps increases our success to quite high levels.

The nature of being human will still throw up complexities for us to solve even when an image credibility framework is achieved, as my neural network analysis of our eye gaze patterns demonstrates. This research demonstrates that physiological signals (eye gaze fixations and saccades) can predict if a person is looking at manipulated or unmanipulated photographs, even when s/he gets it wrong. Further, this type of analysis can also predict the participants' conscious decision (right or wrong), implying that we don't just 'look and decide,' we also 'decide and look.'

These finding are discussed in further detail below.

5.1 Attitudes towards image manipulation

A characteristic of participant attitudes towards image manipulation is that, although homogenous on the whole, the opinions were heterogeneous by group. For example, as illustrated in Figure 5-1 in response to Q4b of the survey the group means ranged from 4 (< somewhat concerned) to 8.8 (very concerned).

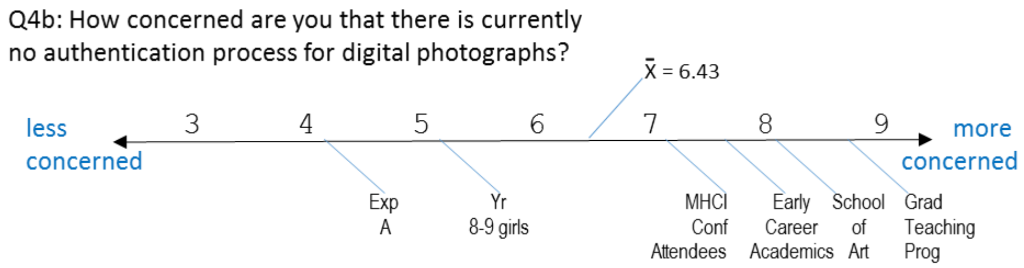


Figure 5-1: Diverse group opinions about the lack of image authentication

This phenomenon would be worth exploring further. However, if taken on average, survey respondents stated that they find it difficult to tell if a digital photograph is manipulated, and after being presented with information about and examples of manipulated images their perception of that difficulty increased. On all three questions testing the attitude of survey respondents to the concerns about image manipulation, the average scores were high, by which we can infer that the 107 respondents felt that the issue of image credibility was an unsolved problem.

5.2 Human accuracy in identifying image manipulations

5.2.1 Identifying if an image was changed

The overall ability of participants to state whether images presented to them were manipulated or not was poor:

Experiment A: 37.5% (SD 0.419) mean accuracy across set 1 of 36 image views with no assistance.

Experiment B: 56.0% mean accuracy across 1,071 image views with pre-training using tested text description. Note that participants had greater success correctly identifying images as unmanipulated (61.3%) than correctly identifying images as manipulated (50.1%).

In the case of experiment A, participants were provided with further assistance from .msci elements in the form of comparison images and image differences map. In the case of these two presentations the success rate climbed to 85.4% (SD 0.297) and 97.2% (SD 0.077) respectively. It seems odd that with the presentation of a difference map that specifically denotes changes, the success rate does not reach 100%, but this may be due to a lack of understanding on the part of some participants about the meaning of a difference map, even though this was explained prior to the presentation of the first difference map. This may warrant further investigation.

In the case of experiment B, participants were familiarised with principles of digital photograph manipulation through an online textual explanation of manipulation techniques including copy/move, splicing, and global transformation, techniques used in the experiment and given quizzes to focus participants' attention on the content. This may be the reason the overall experiment success rate was higher than experiment B.

In addition to low accuracy levels, participants also had low confidence in their decisions. This lack of confidence is revealed in Experiment B, in which grounded theory analysis of participant comments demonstrated that a strong element of uncertainty existed in subjects' decision-making processes with overt verbal statements of uncertainty expressed in 21.2% of all image views (see Table 3-10). One might expect that uncertainty also applied to additional image views, but not vocalized.

This uncertainty may be a factor in the fact that overall, participants in Experiment A responded when asked that that when viewing images they looked first for aesthetics, then meaning, then representations of reality. These three perspectives were ranked in descending order of equivocality; a person viewing an image can quickly determine if they find it aesthetically pleasing, then to a lesser degree understand the meaning of the image, and to a still lesser degree decide its veracity.

However, despite expression of uncertainty in their decisions, participants were still more likely to be confident that they were accurate than they actually were.

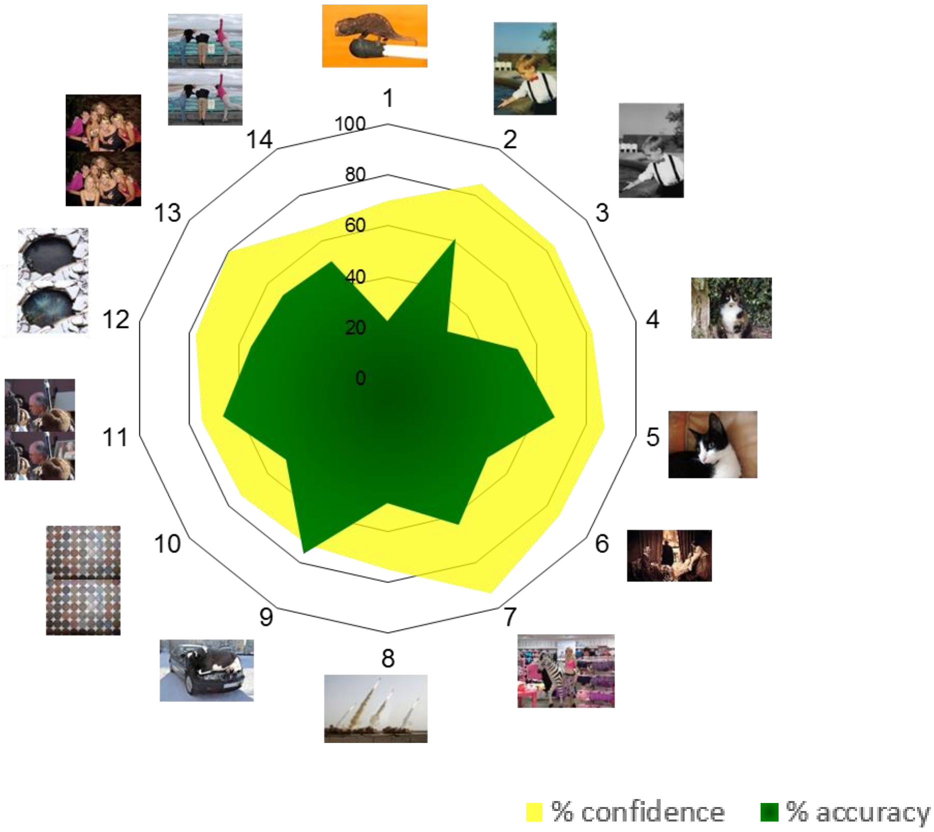


Figure 5-2: Participants uncertain, still more confident than accurate

This implies that we believe we understand the veracity of the images we see daily better than we actually do.

5.2.2 Identifying *what* was changed in an image

It is one thing to identify whether an image has been manipulated or not, and another to identify what had been changed. In both experiment A and B, participants were asked to identify what had been changed if they identified an image as manipulated (whether or not the image actually was manipulated). In experiment B, participants were not offered comparison images and difference maps, so they had to attempt to identify what had been changed without any assistance. The success rate for correctly identifying the nature of the change was low, with only 132 correct identifications of what was manipulated out of 477 valid views of manipulated images, or 27.7% in comparison to 56.0% mean accuracy rate based on 600 accurate identifications out of a total of 1,071 valid image views). Even in experiment A where participants were given a comparison original image, despite participants' success rate at identifying manipulations more than doubling to 85.4%, they often either could not say what had been changed or else misreported or under reported what had been changed. Eye gaze data indicated that participants' gaze in general traversed the two images to identify and compare areas of actual difference.

Table 3-5 of experiment A demonstrated that subjects had looked directly at the manipulated areas with greater intensity than would be predicted by the area of the manipulated regions of the image, although with the low success rate of 37.5% this did not seem to translate into identification of manipulations. While experiment A was not included in NN analyses, this phenomenon prompted my choice of feature set used in NN analyses of experiment B, which demonstrated that the NN was able to use the difference between eye gaze fixations and durations in manipulated regions to identify whether an image was manipulated or not, and whether a participant would say that it was manipulated or not. It would be useful to extend the NN analysis to include these features as they apply to experiment A as well.

5.3 Value of preparedness and training

Participants viewed and judged manipulated and unmanipulated images with different levels of preparedness for the task:

In experiment A, there were 3 different types of preparedness for assessing the manipulation state of

- no assistance
- comparison with original
- comparison with original + difference map

In experiment B, preparedness was uniform for all participants, consisting of including textual information about image credibility in the experiment with a quiz following immediately after the text was presented and immediately before the images were presented.

These varying levels of preparedness yielded the following success rates:

- no assistance (experiment A) – 37.5%
- familiarization with image credibility concepts via text and quizzes (experiment B) – 56.0%
- comparison with original (experiment A) – 85.4%
- original with comparison + difference map (experiment A) – 97.2%

It is important to remember that most people viewing images do not have someone nearby asking them specifically whether an image was manipulated; it is likely that most of the time viewers simply absorb the presented meaning of the image uncritically.

These results highlight the inadequacy of assessment of images through simply looking at them, that is to say with no additional assistance in the form of comparison images or difference maps, and the still restricted ability to assess images even with pre-training.

However, clearly the participants in Experiment B were advantaged over the unsupported participants in Experiment A by virtue of having been pre-familiarised in image credibility issues. This suggests that training in visual literacy would be advantageous for the wider population.

5.4 Neural net accuracy in identifying image manipulations and participant votes

This research was able to identify that in some cases it is possible to show that eye gaze patterns of a group of experiment participants can be classified into those who looked at the manipulated version of an image and those who looked at the unmanipulated version. Some images were more receptive to this type of classification than others. For example, the NN analysis achieved the highest level of accuracy with images 11 and 12 of experiment B and the lowest levels with 10 and 14. There was a marked difference between these two pairs of images in that 11 and 12 had large areas of manipulation and 10 and 14 had quite small areas of manipulation.

The NN effectively demonstrated varying levels of accuracy in determining what the participant would choose (53% to 86% accuracy). Almost all of these accuracy levels were considerably higher than chance, and with a significance level = $p<0.01$, the NN arrived at its correct choices by finding meaningful patterns in the data pertaining to participants' eye gaze for each image, and not by chance (to the 1% level).

This capacity on the part of NNs to determine manipulation status and participant choice is very intriguing, and worth pursuing in future work.

5.5 Image manipulation – not a black & white issue

Even in the face of research such as this that takes empirical steps into the vast ocean of image credibility, the role of photography as evidencing the truth is not as straightforward as it seems. Refutations of the idea of representative photography may be boiled down to four main concerns:

- photos have always been manipulated,
- any photograph is subjective from the moment the photographer frames it up in a viewfinder,
- it's all subjective, and
- come on, what is reality, anyway?

These arguments start from a basic premise: that a photo is *either* real or fake. And it seems debaters believe that if they can muddy these distinct binary waters, we must dispense with the idea of truth in photography.

But it is not logical that a photograph must be assigned to one or the other of these two categories, real or fake. In fact, the most authentic of photographs are usually just a little contrived, and the most manipulated photographs still contain a modicum of reality. As a result, the veracity of a photograph is more a point on a continuum between representative photography, or science, and interpretive photography, or art, as previously described (Figure 5-3).

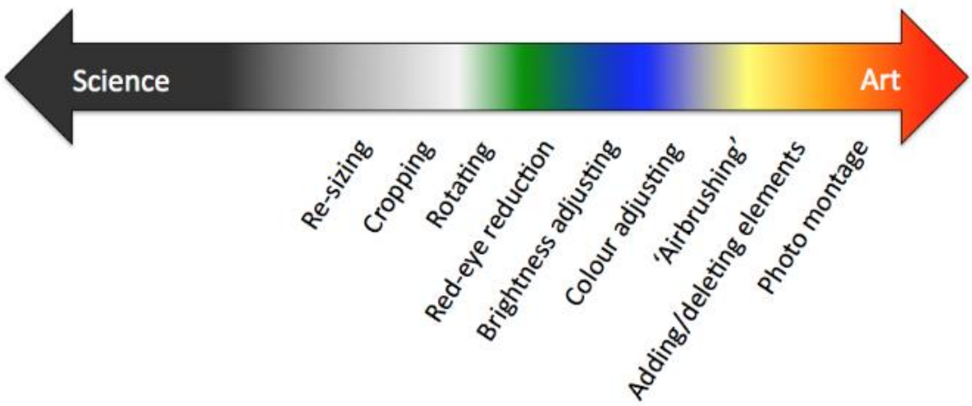


Figure 5-3: Suggested continuum of photo alteration on a Science vs Art scale

Seldom is a photograph a true/false proposition. However, let us take each of these four arguments against photography as a representative medium one by one.

5.5.1 The "*Reality, what a concept"* argument

Firstly let us dispense with the question of what is reality anyway, the argument I like to refer to as the "Reality, what a concept" argument, the moniker for which I took from an album by Robin Williams of the same name (Williams, 1979).

While it is possible to agree that we live in a world governed by the laws of physics, in which a tiny amount of matter in the form of atoms is held in position by a range of forces, and that we inhabit a particular point in space/time, few of us can live our lives that way.

We all choose to believe in a world of family and home, time passing, seasons, stars, adventures and the cycle of life. These are the real things in our world that we can see and touch and feel and ponder every day. Most of us living on Earth choose to believe in this real world and that it matters that we explore and understand this real world as well as our place in it. Photography is one of the ways we do this.

For the physicists amongst us, photography is a record of the visible light photons reflecting from the world around us and impacting on the electrons of our camera sensors. That is well within the level of reality most people can live with. In this level of reality, it *is* possible to speak of photographs as representative of the real world.

5.5.2 The "*Photography is an art, not a science*" argument

As can be seen by the comments of survey respondents, there are varied ideas in the public about the nature and role of photographs. People clearly state in their question responses that images are important, but are not clear on how.

Interestingly, the perspectives of photographers who are entrenched in the discipline of photography (as opposed to casual and amateur photographers as we have all become) is well worth noting. In the snap survey of attitudes towards photo credibility, members of the ANU Department of Photography became *more* dubious about the importance of credibility in photography as opposed to all other groups. This may relate to the malleable nature of photography that can fit into artistic intents as easily as reportage purposes, and the desire of professional photo artists to insert their artistic perspectives into photographs.

This artistic perspective is well illustrated in the following exchange between art critic, Michael Fried (Why Photography Matters as Art as Never Before) and photographer James Welling:

> *"JW: … wouldn't it be interesting if in the future digital era, the idea of photography being an "index" or a trace of light turned out to not be such a big thing?*
>
> *MW: It would be interesting. But what isn't clear is what the very concept of photography and the photographic might mean under those conditions. It might be the evolution of a set of technological procedures for producing large, flat, "depictive" artifacts to look at. But would viewers come to feel that those artifacts weren't photographs anymore, and if not would that matter to them? I have no idea.*
>
> *JW: The question goes back to what is photography? Is it a set of initial moves played out in the early part of the nineteenth century? Or does photography have fewer anchors than we think?"(Welling, 2009)*

These questions are certainly interesting from a perspective of photography as art, but there is no question that photography as reportage, images as data, will always be important and while the line between photography as art and photography is likely to remain blurry, clear differences will always exist.

a  b 

Figure 5-4: Photo viewed by Kirsty Darlaston (a) her embroidery of her eye gaze track (b)

Perhaps it is better to lift our perspective out of an argument about whether photography is a scientific representation of people, places and events into a broader view. In this broader view, perhaps we could acknowledge that when photography came into being, it also spawned new photoARTistic media in which photography plays a supporting role. As an example of peoples' predilection for art, one participant in the eye gaze experimentation undertaken in this research turned her gaze trace created while looking at a photo into a range of artistic outputs, including an embroidery (Figure 5-4).

Landmarks in the 'Photography as science vs art' debate

1839	Daguerre credited with inventing silver halide photography
1840	Edgar Allen Poe: photography is "the most extraordinary triumph of modern science"
1857-9	England's Rejlander and Robinson use new 'science' of photography to create fictional images
1859	Charles Baudelaire declares photography to be a mechanical means of recording reality
1880	Peter Henry Robinson states photography is art.
1890	Peter Henry Robinson states photography is science.
1890s	Jacob Riis and Lewis Hine photograph New York slums and child labour highlighting photography's role of 'reportage'
1902	Stieglitz becomes an activist for photography as an artistic medium and incites the Photo Secessionist movement
1960s	Andy Warhol highlights the art in advertising photography
1978	Susan Sontag states that 'a photograph passes for incontrovertible proof that something exists, or did exist, which is like what's in the picture.
1981	Roland Barthes concurs with Sontag that photography records what is real

Figure 5-5: Landmarks in 'Photography as science or art'

The question of photography as 'art vs science' is an argument as old as photography itself. On this topic, many prominent voices have pondered, including Charles Baudelaire (1859), Stieglitz (1900), Susan Sontag (1978) and Roland Barthes (1981). In 1840, the year following photography's simultaneous but independent invention by William Henry Fox Talbot and Louis Jacques Mande Daguerre, Edgar Allen Poe declared it "the most extraordinary triumph of modern science." In 1859 … Charles Baudelaire emphatically declared photography to be a mechanical means of recording reality, a "very humble handmaid" of art and sciences, “the secretary and record-keeper of whomsoever needs absolute material accuracy for professional reasons.”

In the 1880s, Peter Henry Robinson notably came out in favour of photography as an art, but reversed his opinion based on the cracking of the chemical process of photographic exposure in 1890, which temporarily seemed to leave no room for artistic interpretation. Contemporaneously, Alfred Stieglitz bolstered the perception of photography as an artistic medium … inciting the Photo-Secessionist movement.’ (Caldwell, 2008)

This back and forth continues to this day. But this either/or approach is too narrow for my liking. I believe that photography is both an art and a science; it is erroneous to assign photographs to a binary true/false system when in fact any given photo contains a proportion of truth and a proportion of artistic license, depending on the actions and intents of the photographer. I am not alone in this thinking. In 1978, despite opining that photographs always contain some art, Susan Sontag also stated in her famous treatise *On Photography* that “a photograph passes for incontrovertible proof that something exists, or did exist, which is like what’s in the picture” (Sontag, 1978).

But whatever the proportions of art and science, there is always a degree of reference to the real world in any photograph, and in fact photographers usually depend on it to provide relevance and meaning in their photographs.

So the answer to this key question in photography, is it science or art, is simple in the end. Like the question of nature versus nurture, the answer is both. Sometimes more science than art, sometimes more art than science, and sometimes almost indistinguishably matched. Each photograph must be assessed on its own merits, and this is why understanding what in the photograph is data/reportage and what is artistic interpretation is crucial to our understanding of the photograph.

The inability to reach a consensus over 17 or so decades is symptomatic of the fact that photography is not either science or art, but an admixture of both. It is both subjective and objective, on a continuum with science at one end and art at the other. It is only recently when photo manipulation came into the hands of the many rather than the few that the balance has increasingly tipped away from science and towards art.

However, now there are clear signs that the problems with image manipulation have surfaced in a range of disciplines and are being addressed with a range of tactics of various efficacies. In addition to the continuum of science to art, there is another dimension of image manipulation, and that is whether the photo art is meant to interpret the world in a purely artistic way, or to deceive (Figure 5-6).

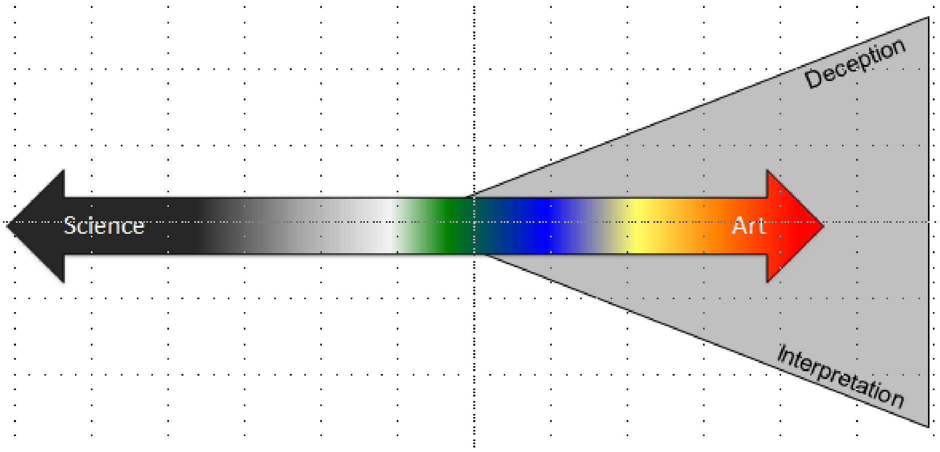


Figure 5-6: Interpretive vs deceptive motivations in photographs as art

Issues with image manipulation often go unremarked in the general populace. The public can simply have not thought about it. For example surveys undertaken in this research (Snap surveys 3.1), in which a short set of four questions about digital photograph manipulation were asked pre and post presentation, in many cases individual participant attitudes towards credibility changed upon being educated in digital photograph credibility.

This failure to consider the implications of digital image manipulation leads to conflicting versions of events, such as the cloning of Lamar Odom into the Kardashian Christmas photo while at the same time the initiator of this minor deception proudly touts that "we always spend Christmas together." Without the preservation of this knowledge about the spliced photograph, future generations of the family could easily assume that the family was physically together at the time of the photograph.

Figure 5-7: Lamar Odom spliced into Kardashian 2010 Christmas card photo. Lamar Odom (top right) was photoshopped in because he couldn't be there.

While a Christmas photograph may seem trivial, it is well to keep in mind that we cannot know all the contexts from which a photograph may be viewed in future. Small features may be clues to answers for questions that haven't been asked yet.

It is important to determine if and to what degree humans are able to perceive image manipulations, and what they think about such manipulations, and these investigations were undertaken in this research.

5.5.3 The "*It's all subjective anyway*" argument

Thirdly, participants argued that photographs as subjective. Of course they can be and often are. As an example, in the photo at Figure 5-8, I chose my camera angle to emphasise the well-deservedness of the farmer's rest: the cattle are secure and relaxed in their pavilion, the hay is fresh and clean, and the young grazier has clearly been working very hard to make it so and has earned his nap.

Figure 5-8: Australian Lowline Cattle and grazier, Royal Canberra Show 2006

As photographers we impact upon our photographs by our choices of subjects, the angles and camera settings we choose, and how we present our images (online, in a photobook, as a print, etc.).

But regardless of these subjective subtleties, when all is said and done, a photograph is still a specific and accurate reflection of the scene in front of the lens at the time the shutter is released and the light floods the camera sensor. At that moment the photo is a real image of the real world. Undeniably, the young farmer, pavilion, cattle, and hay were present as photographically described when I stood there and snapped their photo.

Subjectivity does not equate to subject invalidity.

5.5.4 The "*Photo manipulation is nothing new*" argument

And lastly, there is their argument that photos have always been manipulated and there is nothing new about it. I am often told that truth in photography has always been elusive, with photographers working in darkrooms through the ages creating any number of creative but false images, which by virtue of association with all other contemporaneous photographs, negates the possibility of photos ever being regarded as true representations of the real world. I doubt the World War II soldiers who went to war with a photo of their mum or their sweetheart in their breast pocket would agree.

Or tell former Chief White House photographer Eric Draper who described his job as "to create a visual archive of official and non-official events of the President's time in office" (B. Baldwin, 2016) that his photos are negated by such an argument.

Figure 5-9 Broad billed Hummingbird, Kenton C. Lint Hummingbird Aviary, San Diego Zoo, 2005

Further, there is much that *is* new about photo manipulation. Digital photography and the availability of the Internet exponentially exacerbate the problem of photo manipulation. It was once the case that few photographers could tinker with their photos, now almost anyone can. It was also once the case that photos were physical objects that tended to remain within a relatively small circle of family and friends who knew all the people and places in the photos (or knew someone who did); now photos are electronic constructs that whip around the world at the speed of light with little context or explanation.

I took the photo at Figure 5-9 and have done nothing to it other than resizing it. The bill of the hummingbird below is orange, it was not 'painted' that way. Its feathers really are a jewel-like turquoise, cobalt and emerald. All of these authentic details allow us to know that it is a Broad billed hummingbird.

We humans have generated an incredibly large pool of photographs, particularly over the past 15 years or so since digital photograph was introduced. It is important that we consider the benefits of instilling and bolstering the security of the meanings in this amazing resource; it is a reservoir of images to inform our understanding of society, history, nature and a plethora of disciplines for archives, galleries, and knowledge production, as long as we know what we are looking at.

5.6 A symbolic formula for calculating image credibility

When the relationships between the various elements affecting image credibility are understood, it is possible to posit a theoretical formula to quantify the relationship between the presented image and the original photograph. Of course it would be quite difficult to establish reliable figures to plug into this calculation. At present such figures would be as subjective as the opinions of the distributors and consumers of the images themselves. However there is the possibility that in future the formula can be populated with at least some real numbers; recently such quantification was attempted in respect of image manipulation of models to create a meaningful metric of photo retouching based on geometric and photometric changes (Kee & Farid, 2011).

For the present, it may be useful to consider these relationships in a formulaic manner. Such a theoretical formula based on the descriptions of the elements affecting the veracity of images would therefore be:

$$\frac{\text{Presentation image}}{\text{Original image}} = \frac{1 - \sum \ell, g, a}{1}$$

Where: ℓ = local manipulations,
g = global manipulations
a = artifacts such as compression, display resolution

This formula can be further expanded when the elements of Metadata, Context, and Photographer's Reputation are taken into consideration. If we assume that the metadata of the image has not been tampered with, then metadata will always add to the credibility of an image. Context and Photographer's Reputation are elements that may add or detract from the credibility and authenticity of the image. Thus

$$\frac{\text{Presentation image}}{\text{Original image}} = \frac{(1 - \sum \ell, g, a)}{1} + \mathcal{M} \pm \mathcal{C} \pm \mathcal{P}$$

Where: ℓ = local manipulations,
g = global manipulations
a = artifacts such as compression, display resolution

$\mathcal{M}$ = in camera metadata and user-entered metadata
$\mathcal{C}$ = context in which the image is presented
$\mathcal{P}$ = photographer's reputation

$\mathcal{M}$ is the metadata of the image provided by technological means (metadata recorded by the camera at the time of the photograph). Metadata, assuming it is untampered itself, will always add to the authenticity of the image by providing a range of important data about the image capture such as date, time, location, camera details etc.

$\mathcal{C}$ is a quantity that can be a positive number, thus increasing image credibility, if the context in which an image is presented is supportive of the truth of the image, in which case the value of the Presentation image over the Original image could be considerably higher. Equally, it could be a negative figure, if the contextual elements within which an image is presented are false or misleading, in which case the value of the Presentation image over the Original image could be a negative figure, indicating that the Presentation image is worth less in representing reality than the Original image.

$\mathcal{P}$ Like Context, a Photographer's reputation can be additive or subtractive. A photographer with a known reputation for manipulation of images should be considered more likely to manipulate the image under examination, and by the same token an image by a photographer with a reputation for not manipulating images or clearly describing any manipulations can be viewed with more surety of credibility.

Note that the idea of an image being created using staging, like the Cottingley fairies, and thus misrepresenting reality is quite relevant to the credibility of the image, however it is outside the scope of this formula for two reasons. First, other than through notation in metadata or context, there is no way to incorporate the alterations made to the actual scene being photographed. The baseline for the original photo (and thus the formula) is set at the control point of the image being recorded by the camera sensor. Further, one could argue that even though the scene was staged, it is in fact the real scene recorded by the sensor of the camera, staging and all.

This formula also does not take into account the way in which we as humans perceive either the original or the presentation images, which as we have seen is an important second part of the two way communication of information that is representative photography.

5.7 Proposed image authentication framework

It is clear from the poor ability of humans to recognise manipulations in the images we consume that it isn't enough to have specialist techniques to identify falsifications. We need to be able to understand them at a rate more immediate, and more relevant to the speed at which we consume them.

And then there is a larger question – what do these photographs mean? Once they are removed from the stakeholders, they begin to rapidly lose their meaning. The names of the subjects, places, events are lost. The knowledge of what was changed in the photograph is lost. The relationships within a group of associated photographs are lost. We need a technology that captures these elements that showcases what technology can do to be bigger and better than the problem it faces.

At the moment, it seems like the credibility of photographs is something of the past, because there is so much indiscriminate changing of today's digital photographs. But we have the ability to preserve so much more of the meaning of a photograph, and at the same time restore the credibility of digital photographs by capturing an unchanging version of the original photograph, like a fly in amber.

Assessing the authenticity of images can take a proactive and targeted approach to comprehensively preserve our understanding of the images we produce and consume and the meanings they bear. My research focussed on this proactive approach: an *a priori* approach in which future photographs can be enclosed in a file format allowing photographers to a) safeguard the credibility of their original image, perhaps out of camera, b) provide manipulated presentation versions of their images separate to the original, and c) display a range of technical, contextual and metadata information about the image.

Rather than retroactively determining if an image is credible, photographers can proactively assert that they are. The framework by which this optimal solution can be delivered is a technological and social solution.

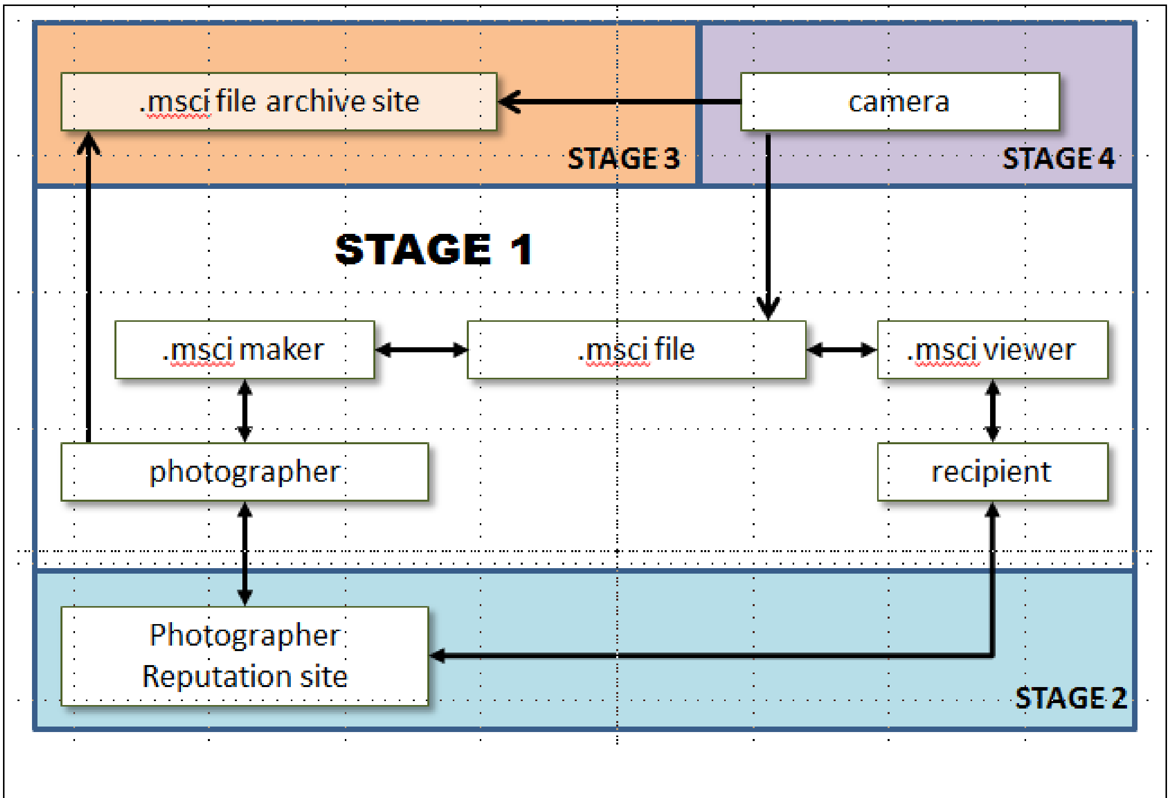


Figure 5-10: Stages image authentication framework

The solution proposed to address the issue of image authentication by developing a technical and procedural framework that fuses technology and social solutions.

5.7.1 Stage 1

At the heart of the solution is the .msci file format which presents digital images and associated information in a standalone format. Technical aspects that are new in this format are:

- the idea of ‘packaging’ original (proof) and modified (presentation) versions of the image.
- extraction of metadata from the obscure locations in which it currently resides to a clearly visible display.

The specification details a new file format that provides an authentication and presentation framework for original and manipulated photographs and their associated information. The name of the format is tentatively called .msci, an acronym for mobile self-contained image.

The format of the .msci file allows photographs to be distributed in such a way that the consumer viewing the image is able to easily see:

I. The image the distributor has derived from the original photograph incorporating various photoprocessing ranging from red-eye reduction to montage. This is known as the Presentation.

II. The original photograph(s). This is known as the Proof.

III. Automatically captured associated information including the photograph metadata such as aperture, shutter speed, white balance, etc. and additional information provided by the distributor, such as organisation logos, standards logos, narratives, names of subjects, GIS coordinates, etc. This is known as the Provenance.

Thus, a .msci file can be graphically represented as trio of layers: Presentation, Proof and Provenance. In theory there may be between 0 and ∞ of each layer; in practice there would normally be at least 1 each of the Provenance and Proof layers, up to a size that is feasible for transmission between the distributor and the consumer.

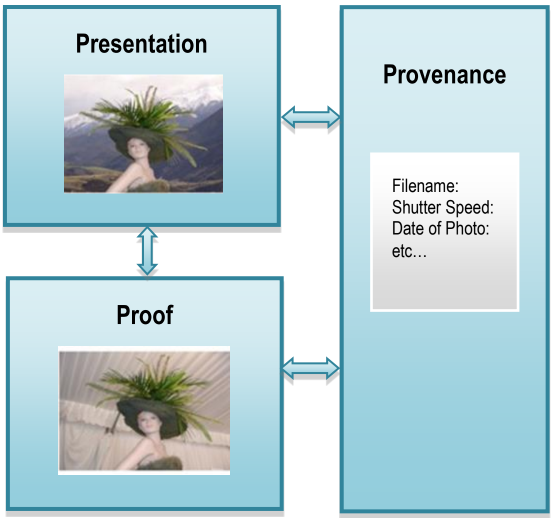


Figure 5-11: The 3 Ps - Presentation, Proof and Provenance

Using this format an image can be transmitted as a self-contained package of information:

Presentation(s)
- Can easily be compared to proof layers
- Photo artists can show their sources
- Photo manipulations normally become evident

Proof(s)
- No Proof layer highlights 'presence of absence' and photo cannot be authenticated
- Proof(s) can be a copy, authentication string or similar, e.g. accession # of collecting institution

Provenance(s)
- Can contain easily accessible metadata
- Can hold range of contextual information
- Available for both Proof and Presentation images

The format of the file is a container file, which allows for a diverse range of data to be included in a single file. Container files also allow workflow to occur, which could be used for authentication mechanisms produced in camera and/or connected to an external locus on the internet. Such a file in use might look similar to that portrayed in Figure 5-12.

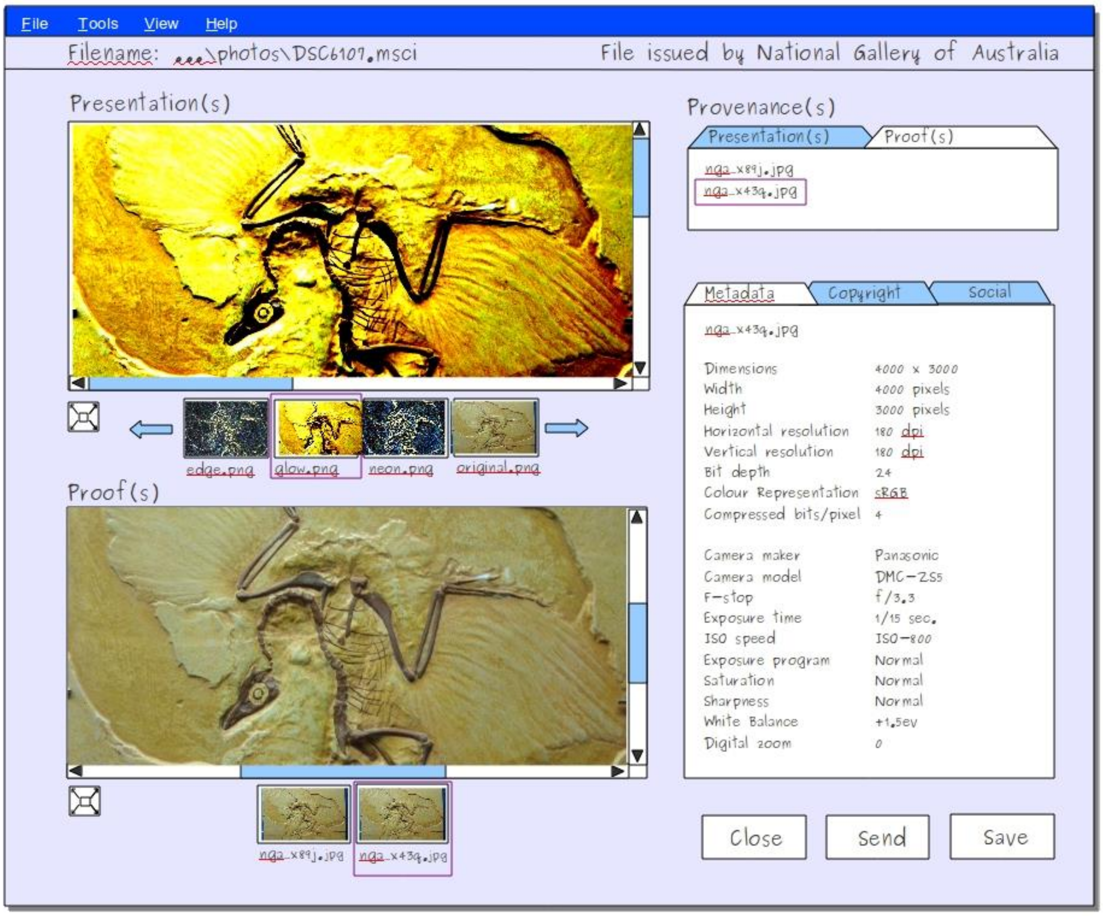


Figure 5-12: Mock-up of .msci file in use

The file format will accommodate backward compatibility (pre-existing images can be ingested into the Presentation and Provenance) and forward compatibility (mobility and self-containment useful as the world increasingly goes 'mobile.')

It is important that this development be done first, in order to create a minimum viable product that will allow engagement with photographers for additional development and progressing to stage 2 in which photographers will be asked to engage with social reputation buiding, of which the .msci files produced are the tangible expression.

5.7.2 Stage 2

The framework does something new in the field of image credibility in that the technical solution is linked to a social solution. The framework harnesses the power of social peer pressure, in that it creates an impetus and capacity for photographers to positively assert that the original image is original and not modified, or, if it is modified, in what way. This aspect engages the photographer with issues of honesty, reputation, artistry, meaning production, history, and many other elements of our social fabric.

While there will always be individuals who break the rules, the majority of producers of digital photographs adhere to common social conventions, and building this social aspect into the framework in connection with the .msci file format offers them a way to demonstrate how their work represents, reflects, or adds to our understanding of the world. For consumers of digital images, this format elevates both the meaning inherent in the image(s), and their ability to have confidence in their relationship to the real world.

5.7.3 Stage 3

Stage 3 and stage 4 may be interchangeable or may take place simultaneously, depending on partnerships and resourcing.

Stage 3 is the point at which a world-class, secure archive for .msci files can be created. It may be a good opportunity to work with an existing archive, such as Getty Images, or a museum such as the Smithsonian or the National Library of Australia, or it might be preferable to create a new archive to allow for clear branding of the archive's purpose and to eliminate the possibility of potential partners' existing agendas confusing the goal of this archive.

5.7.4 Stage 4

Ultimately it would be desirable for .msci files to be produced at the point of the photograph being taken. In order to effect this, a camera manufacturer would be needed. It would be preferable to work with a major camera manufacturer such as Canon or Nikon to assist with widespread adoption of the image credibility framework and .msci file. Partnering with a major camera manufacturer may not be possible until some evidence of the framework achieving traction has been demonstrated. It is also possible that a new prototype camera could be engineered given the right resources.

More details on the proposed image authentication framework can be found in *Appendix B: Design Specifications for Image Authentication Framework including Content and Semantics for .msci file format.*

6 Conclusions and future work

Firstly, I conclude that there is much more work needed in this field. The outputs of my research in many cases only tantalizingly hinted at a much bigger and more complex discipline whose essence is only one of numerous elements that will take many years to define. Perhaps ‘image credibility theory’ is as good way to refer to it as any. But that this field takes in theory from many disciplines - history, sociology, computer science, visual arts, psychology, communications studies, law, physics and others – is clear. What is also clear is that, as many disciplines as may be involved in defining and solving the problem of image credibility, there are many more impacted by it – probably almost any discipline one might name. This makes it a big problem to be solved, even if it is not quite as well recognized now as it will be in future.

Having said that, let us consider what we learned in this piece of that grand puzzle.

6.1 Conclusions

In a cross-disciplinary fashion, I began this thesis with an exploration of the social impacts of image manipulation, and found that seated within the progress of photography has always been a perpetual desire by people to tinker with the images we create even as we create them. We use whatever tools are to hand, and the more effective the tools, the more effective the manipulations. Camera and software technologies have pushed the event horizon of such manipulations back right into the depths of the camera sensor image capture and processing so that many images are now not just ‘born digital’ but ‘born digitally enhanced.’ While at this point in history, the expectation of credibility of images is still high, their actual credibility continues to diminish.

That the problem is being recognized and tardily addressed is evidenced by the fact that new rules detailing what is acceptable in image enhancement are being imposed by high profile photographic associations, while simultaneously, attempts at image security and new secure cameras are being made.

What is also clear is that ability of viewers to discriminate between manipulated and unmanipulated images without additional assistance is poor. Even with overt questioning as to whether or not an image has been manipulated, results without robust assistance were at or below chance:

- no assistance (experiment A) – 37.5%
- familiarization via text and quizzes (experiment B) – 56.0%
- comparison with original (experiment A) – 85.4%
- original with comparison + difference map (experiment A) – 97.2%

However, the snap survey results indicated that people have a reasonable level of awareness of their deficiency in ability to identify the credibility status of the images they view, as well as a concern about it (Snap Survey: Attitudes toward image tampering p. 66). This is further borne out by the grounded theory examination of verbal responses in Experiment B in which the themes of uncertainty and logic played roles in participant voting.

6.1.1 Eye gaze

Based on the experiments undertaken in this research, it seems that being presented with an image is in itself insufficient to reliably identify if the image we view has been altered. We get it wrong 44% (experiment B) - 67% (experiment A) of the time, and we are even poorer (>70% of the time) at telling what has been manipulated in the images we do successfully identify as manipulated. However, eye gaze tracking shows that increased attention (longer and more fixations) to manipulated areas of photographs can help.

Our eye gaze is reflects the features of an image that are non-consciously noted, but this does not necessarily translate to verbal accuracy. We may miss smaller manipulations via the ‘hiding effect’ of larger manipulations. If we lack additional knowledge about the photograph, we may resort to the use of logic in determining the veracity of the image; this is often drawn from personal experience (which may or may not be relevant), and even after deciding we are still uncertain of our conclusions.

The information we receive visually is complex both in the actual data and the way in which we process it to arrive at our determinations. We have seen that neural networks can predict both the nature of the image (manipulated or unmanipulated) as well as our decisions based on our eyegaze, which further demonstrates the importance of visual cues to support our cognition. Any image credibility solution framework needs to incorporate such support.

6.1.2 Human attitudes towards image manipulation

My experiments in human perception of digital images make it clear that we are not good at identifying manipulations in images. With more and more images arriving before us with a range of manipulations ranging from ‘enhancement’ to morphing, one of two outcomes must inevitably result: a) many of us may lose faith in images as representative of real people, places and events, or b) a solution for preserving knowledge about the representative qualities of images will be found and the line between photos as artistic interpretation and reportage will be more clearly delineated.

6.1.3 The key control points for proactively authenticating images

There are two key control points for proactively authenticating images. The first is technological; the central control point for a digital image is the image sensor. The second is human centric: the central control point for creating meaning and credibility in a digital image is the photographer. Given that we know that no system is proof against concerted hacking and will eventually fail, any solution for image authentication needs to incorporate a human element, which brings with it the benefits of reputation and adjutant knowledge of the circumstances of the image.

6.2 The future of photography

Unless there is a global catastrophe of massive proportions, it seems that the future of photography is assured. We humans cannot get enough of images, and we are recording our lives and environments and events in ever increasing numbers. But will the future of photography be framed as it has in the past? Yes and no.

6.2.1 Photography as 'science vs art' debate is still relevant

It is interesting to see that the question of photographs as 'science' or 'art' is still playing out today, even when there has been almost two centuries for this to be settled. To an extent, it is similar to the argument of 'nature' vs 'nurture' in which we continue to debate whether a person's character is formed by genetics or upbringing, even when countless studies have shown that our characters are formed by both of these factors. Not surprisingly, it appears that photographers argue in favour of artistic freedom in photography, and consumers of images as reportage argue for truth in their visual information. It is unlikely that this debate will reach a conclusion satisfactory to all parties, and so this tension between photography as an art or science is likely to remain entrenched in society far into the future, with no resolution in sight other than a guarded truce. However, all of the richness of ways in which we engage with our photographs is scaffolded by the levels of reality and unreality that are expressed in those photographs, and therefore it will always be important to know what these levels are.

6.2.2 New cameras and other image capture technologies

Photography as a discipline remains a dynamic, progressive field moving swiftly into the future. While SLR style cameras can be expected to continue for the foreseeable future, we can expect cameras to continue to diversify and for new uses to be found for them. Some examples (described below) do not even match the current ideas of cameras. Of course phone photography, or phonography or iphonography as it has come to be called (Halpern & Humphreys, 2016), is a good example of cameras being created in a non-traditional form. But there are many others.

Figure 6-1: Tobii Glasses 2 – contains a wide angle HD 1920 x 1080 pixel camera and video at 25 frames per second.

A burgeoning field for photography is the emerging wearable cameras, in the first instance cameras as a photographic interface in glasses, already on the market in the form of Google Glass and Tobii Glasses 2 (Figure 6-1).

In addition, completely new types of cameras are joining the market. One example that confounds the idea of the moment of truth in photography is the Lytro camera, which offers viewers more power to see into the photographs with which they are presented. This new style of camera allows photographers and those who view their photographs to focus on an image *after* the photograph has been taken. At least initially, it will be more difficult to manipulate images taken with this camera, however it is worth noting that the camera is already sold ‘bundled’ with a few basic image manipulation tools.

This power to view images as the user wishes is created by the Lytro camera capturing all the light incident upon the camera sensors and thus offering viewers the ability to change the focal point displayed in the on-screen image, after they have been taken.

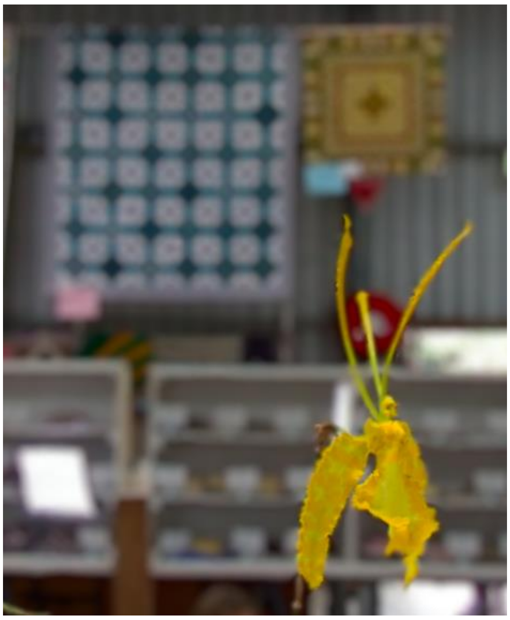
a  b 

Figure 6-2: Examples from the Lytro camera, a ‘post photo capture’ refocusing camera

Regional country fair photo with focus on background quilts (a) and foreground orchid (b) from a single Lytro photograph. (photo by Sabrina Caldwell)

6.2.3 Image management systems

However an image is acquired, it must then be managed. Where will it be stored, who will be allowed to access it, how will it be distributed? The need to develop systems to accommodate the main types of social uses of images (memory, social relationships, self-expression and self-presentation) exists now and will be on-going in future as technological improvements and social trends change (House et al., 2004)

6.3 Future work

As mentioned earlier, there is a lot of work to be done in this discipline. However, confining the list to likely next steps arising out of this research, I believe there is great value in pursuing further research in human perceptions of images, our attitudes toward and understanding of the impacts of image manipulation, and bringing new tools like neural network analysis to bear in examining these issues. Importantly, I will be building on the initial .msci file format work to create the full suite of .msci maker, file format, and .msci viewer as a first step in deploying the full image authentication framework I envisage.

6.3.1 Human perceptions of manipulations in images

Because results in human accuracy and the neural net analyses of their processes in arriving at that accuracy were so variable from person to person, and from image to image, it would be useful to attempt new experiments in which as many variables as possible are controlled. Would results be more uniform if the participants were of similar age, gender, cultural background? Would they be more consistent if the images themselves were sorted into levels of complexity? These would be interesting questions to explore.

In experiment A the presentation of a difference map between original and manipulated versions of images did not yield a 100% success rate, and further investigation as to why would be useful, especially in light of the expectation of using difference maps in the .msci file format at the heart of the authentication framework.

I noted in the results for the snap surveys that some participants believed they were better at spotting fakes after exposure to information and/or examples of them. It would be useful to conduct experimentation to determine if this is true. Also, the diversity of opinions from group to group would be interesting to explore further, as such differences may indicate different motivations, life experiences, conflicting views or other factors.

It would be useful to determine to what extent different types of visual literacy training can support people’s ability to think critically about the images they consume, and this would bear further investigation.

There are still many interesting questions to explore in respect of the qualitative results, such as individual variations in performance. Are there in fact 'super-detectors' amongst the subject population who excelled in identifying whether a photograph had been manipulated or not? What seemed likely to help or hinder people in coming to their decisions? To what extent does pre-familiarisation or visual literacy training help?

Further, it would be useful to simply increase the volume and type of images tested, and increase the use of grounded theory analysis to capture more of the qualitative data arising from experimentation, especially in respect of cognitive and attitudinal perceptions of tampered images.

6.3.2 Neural net analysis

The analyses undertaken in this research were limited to pattern classification in which I learned that, to varying extents, a supervised NN is able to learn the difference in eye gaze features depending on image manipulation state and participant determinations.

It would be of value to analyse the eye gaze data across various images in varying combinations of number and type of manipulated and unmanipulated images using clustering to see if it is possible to separate out eye gaze tracks into various groups corresponding to those combinations. It would also be useful to attempt to identify if individuals or groups of individuals use particular eye gaze strategies when viewing manipulated images, and how those eye gaze strategies altered in response to information provided or questions asked about the images. Some specific questions to be explored are:

- What characteristics of manipulations cause classifiable differences in eye gaze?
- Can eye gaze pattern information be used to predict whether an image has been manipulated? Can it be used to interrogate historical images to determine the likelihood that they were manipulated, and how?
- What is the difference between the two categories of eye gaze relating to participant choice?
- Why are the predictive abilities of the NN so different for different images?
- Why is there a bias one way or another towards participant choice in different images?

- Is there a correlation between eye gaze strategies identified by an NN and:
  - the themes of uncertainty and logic noted in Experiment B?
  - the attitudes towards image manipulation noted in Experiment A participants?
  - the quiz results of Experiment B?

While experiment A was not included in NN analyses, it was the fact that participants eye gaze fixations and durations appeared to linger on manipulated areas that prompted the choice of feature set used in NN analyses of experiment B, which demonstrated that the NN was able to use the difference between eye gaze fixations and durations in manipulated regions to identify whether an image was manipulated or not, and whether a participant would say that it was manipulated or not. It would be useful to apply NN analysis to include these features as they apply to experiment A as well.

There are many other types of images that can be usefully investigated with NN and other analyses, such as the post-capture photographs discussed in 6.2.2, and novelty images like Magic Eye stereo optical images in which specialised viewing skills are required.

NN testing and analysis would also be a very useful element of user testing of the .msci format element of the authentication framework, to identify what features of the framework are best utilised by viewers performing tasks with the .msci format tools.

6.3.3 Development of full image authenticity framework

As mentioned in the introduction, at the moment the social recognition of the import of digital image credibility problems is still quite low, which means that photographs as evidence of the real world are still valued, despite the proliferation of manipulated images. The bias to believe in photographs also still seems to exist judging from the tendency for experiment participants to choose unmanipulated when making a choice. This allows for authentication solutions to be introduced while photographs are still widely regarded as accurate representations of actual people and places.

There is further experimentation that needs to be undertaken in respect of this framework that was simply outside the scope of what I could achieve in the term of this research. In particular, I would like to do further experiments designed to understand how metadata impacts on our understanding of images, especially metadata describing image manipulations. I would like to conduct further tests of the role of comparison images and difference maps (which proved to be the most effective support for people attempting to identify manipulated images), and incorporate qualitative analysis that gleans further insights into how people would use and feel about such an image format. And additionally, I would like to determine how social pressure such as reputation will affect people's perceptions of images and the photographers who produce them.

Many of the steps required in developing my image authentication framework are achievable without having a camera manufacturer as a solution partner. However this would be highly beneficial.

6.3.4 Implementing next steps

The research activities I outlined above, despite being constrained in scope from the perspective of the field of image credibility theory, are still very large and complex research activities. They are also cross-disciplinary. I have discovered while doing the current research that cross-disciplinarity adds a further effort impost by widening the field of literature to be consulted, increasing analysis time by demanding both qualitative and quantitative analysis, and additional time interpreting and synthesizing the outcomes of experiments.

Accordingly, the next steps in progressing this important research should include attracting additional researchers to my project from both computer science and other relevant disciplines such as psychology, sociology, the sciences, and the arts. Within a university environment this can commence with higher degree research students, and hopefully expand by applying for grant funding to allow longer term resourcing. There is definitely a place for industry in this research, especially camera manufacturers as well as peak bodies concerned with the short and long term implications of untrustworthy images illustrating our science and history and telling our human story.

6.4 In summary

This thesis has been a journey of discovery in which I have had the opportunity to explore the entwined and entrenched nature of image tampering within the store of photographs and other images that defines our past, present and inevitably our future. I have been able to investigate this phenomenon not only from a technical perspective, but also from a human perspective. I have learned more about the landscape of this problem, and hopefully promulgated awareness of the need to preserve the credibility of our images to listening ears.

I believe the future holds an ever greater demand from society for increased image security and preservation of the meanings inherent in the images we produce so profligately today. With further research, and development of frameworks similar to the one I have outlined, we may just be able to address this demand effectively. Ultimately future photographs may not just be 'born digital,' but also 'born credible' – and stay that way.

7 Glossary

CCD – Charge Coupled Device

CMOS - Complementary-symmetry metal–oxide–semiconductor

CRC – Cyclic Redundancy Check

HMAC – Hash-based Message Authentication Code

Image – Generic term denoting a visual depiction, encompasses photographs, manipulated photographs, photo art, data visualisation, charts and graphs, etc. For the purposes of this thesis, the term ‘image’ refers largely to the set photographs, manipulated photographs, and art.

MAC – Message Authentication Code

Metadata – Information about the image, such as EXIF data

MFile (.msci term) – convenient name to call the .msci file format

MPack (.msci term) – convenient name to call the .msci packaging application

MView (.msci term) – convenient name to call the .msci viewer

Photo art – An image that incorporates photographs or parts of photographs but that also has manipulated components or non-photographic components included

Photograph – An image that has been created by using a camera sensor recording the light reflected from a real world scene

Presentation (.msci label) – End version of the original photograph

Proof (.msci label)– Original photograph

Provenance (.msci label) – Visible listing of metadata provided in-window

SHA-1 – Secure Hash Algorithm 1, produces hash digests of 160 bits

SHA-2 - Secure Hash Algorithm 2, produces hash digests that are either 224, 256, 384 or 512 bits

SLR – Single Lens Reflex

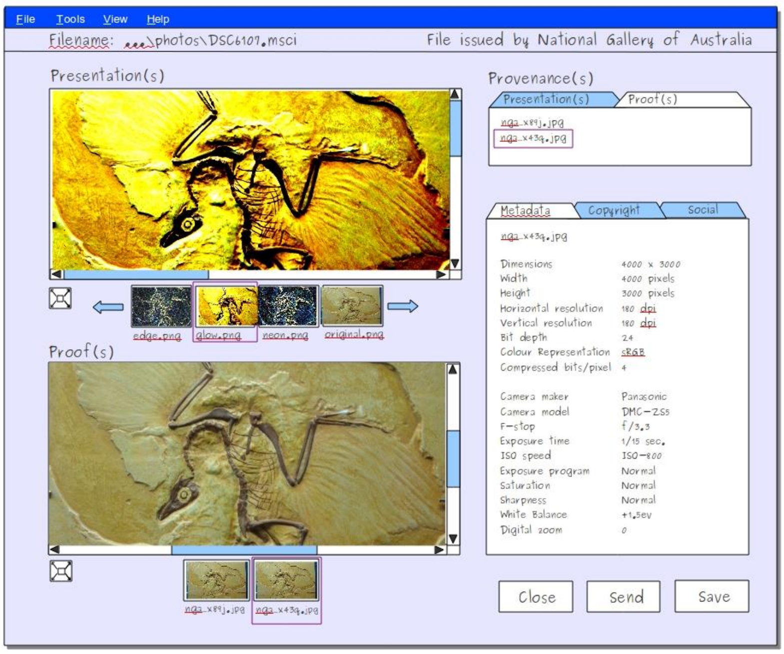
8 Appendices

8.1 Appendix A: Index of photo hoaxes noted in this document

8.2 Appendix B: Design Specifications for Image Authentication Framework including Content and Semantics for .msci file format

Design Specifications for Image Authentication Framework including Content and Semantics for .msci file format

Sabrina Caldwell
November 2016



Sample of .msci file format as it might be used

Contents

List of Figures

9 Introduction

As time goes on, understanding the extent to which a digital image can be believed is becoming increasingly important as ‘enhanced’ digital images proliferate, and purport to represent the reality of individuals and our environments, record significant human events, and are used as supportive evidence in areas as diverse as academic research and law enforcement.

Many strategies and technical solutions for examining ‘sourceless’ images for manipulations exist, but there are currently no holistic solutions for securing the credibility and source information of images into the future. This is despite the fact that image consumers are concerned about the credibility of their images and many image producers desire the ability to assert the veracity of their images.[4]

The proposed image authentication framework is designed to address this problem by combining a technological solution based on a mobile self-contained image file format with an online verification and social solution.

Current approaches to credibility in digital photographs can be divided into two types: digital forensics and digital authentication. Digital forensics are a collection of techniques that interrogate an existing image and identify a range of data artefacts arising from the photo manipulation process. The outcome of this interrogation is either a positive statement that the photograph has been manipulated (and how) or an indeterminate outcome in which it is not known if the image has been manipulated. These techniques, while available to specialist organisations such as the police and other security organisations, are of limited availability to general consumers of photographic images. Digital authentication is a proactive process in which the image file is captured and secured against modification. Although several camera manufacturers have attempted this (including Canon, Minolta, Nikon and Kodak), to date no solution has been successful, all having been ‘hacked’.[5]

---

[4] Caldwell phD Thesis research

[5] The predominant hacker of these solutions is the Russian organization Elcomsoft.

It is widely concluded that no technological solution will be 'hack-proof'. As a representative of Canon put it, "It is a technical arms race."[6] This suggests that the solution for digital image credibility must incorporate elements outside technology. For this external element, this solution turns to an often neglected key control point of photographic credibility: the photographer.

This image authentication framework widens its view to take in the perspectives and needs of the photographer and the people who view the photographs. A top level view of the framework architecture is provided at Figure 9-1; it is explained in this document.

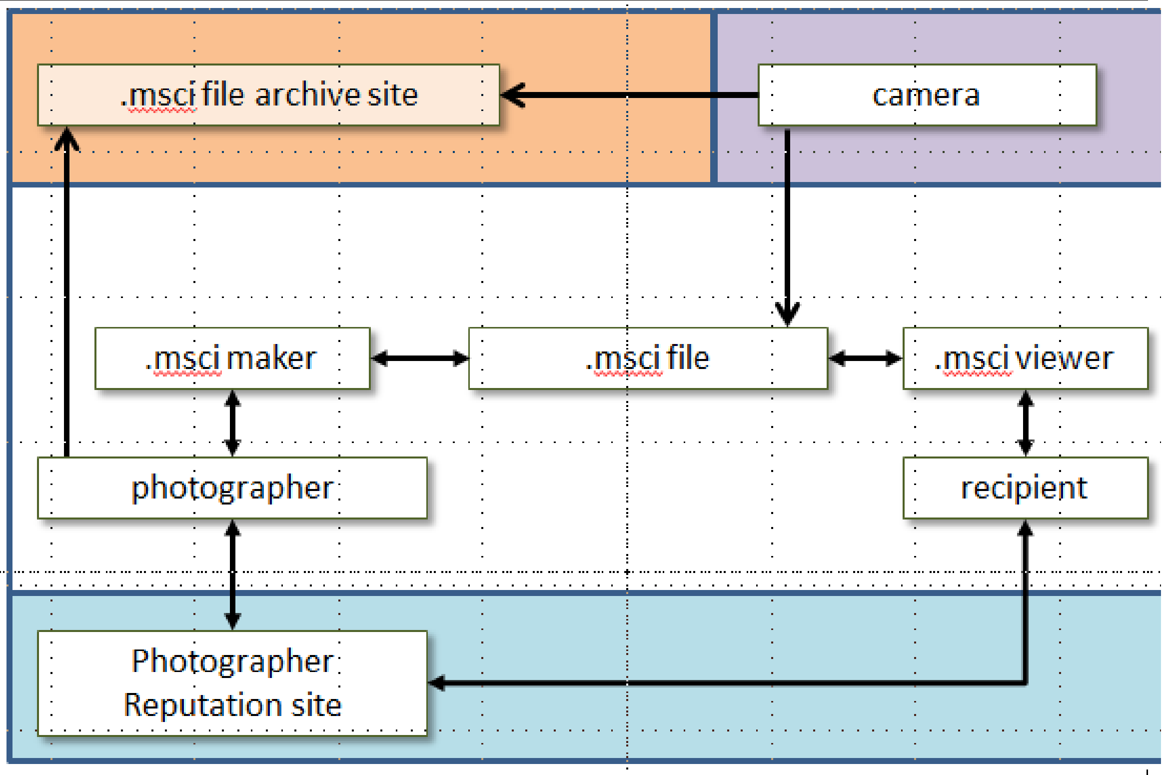


Figure 9-1: Image authentication framework

[6] CISRA (Canon Informatin Systems Research Australia) 'Lunchtime seminar' presentation July 2012

9.1 Purpose

The purpose of this document is to specify the design and technical specifications of the proposed image authentication framework including the core image file format and to identify the relationships between the elements of the framework.

9.2 Scope

The system is intended to be developed in 4 stages Figure 10-1 with the .msci system as stage 1.

This document provides lower level detail in respect of stage 1, with higher level discussion of the characteristics of stages 2-4.

While this document focusses mostly on development, any development incurs costs, and some of the larger scale elements such as the photographer's reputation site and the online .msci archive site may have quickly escalating costs of implementation. Indications of funding approaches are therefore included in this document.

9.3 Intended Audience

In the first instance, the intended audience is restricted to the author, ANU colleagues and thesis examiners. Ultimately however this document is intended to be the foundation for developing the proposed system and at that point the audience will be widened to include other project partners as specified at that time.

9.4 System names and relevant terms

The core image file format currently has the working title of *mobile self-contained image file*, with a potential extension name of .msci.

The overall framework name currently has the working title of *image authentication framework* but it would be anticipated that a socially appealing name would be identified prior to release.

9.4.1 Relevant terms

The following list explains terms used in this document:

Term	Definition
Consumer	Person or organisation who receives and views or otherwise uses the .msci file
Distributor	Person who produces and distributes the .msci, especially the photographer
Essence-only file	This is a single aspect file such as a text file or standard image file and is best understood by comparison to a wrapper file in which several files are enclosed in a single packaged file such as a zip file.
Metadata	Information about the image, such as EXIF data
MFile or .msci	working 'short names' of the mobile self-contained image file format
MPack	working 'short name' of the .msci maker/packaging application
.msci system	The .msci system comprises the MPack, MFile and MView
MView	working 'short name' of the .msci viewer
Presentation	Modified version of the original photograph
Proof	Original photograph
Provenance	Visible listing of metadata provided in-window
Wrapper file	This is a file format in which several files are enclosed in a single packaged file such as a zip file. It is also known as a container file.

9.5 Relevant documents

There are no required documents relevant to this specification. However it should be noted that much of the design of the framework and .msci file format is informed by research undertaken at ANU as part of a PhD program, and therefore the PhD thesis *Perceptions and problems of digital image manipulation and framing digital image authenticity* (2016) by the author and associated published papers contained therein are useful adjuncts to this document.

10 System overview

The solution proposed to address the issue of image authentication is a technical and procedural framework that fuses technology and social solutions.

At the heart of the solution is the .msci file format which presents digital images and associated information in a standalone format. Technical aspects that are new in this format are:

- the idea of 'packaging' original (proof) and modified (presentation) versions of the image.
- extraction of metadata from the obscure locations in which it currently resides to a clearly visible display.
- A difference map to highlight changes.

The framework as a whole also does something new that is not technical but social - it harnesses the power of social mores, in that it creates a positive assertion on the part of the distributor that the original image is original and not modified. This aspect engages the photographer with issues of honesty, reputation, artistry, meaning production, history, and many other elements of our social fabric. While there will always be individuals who break the rules, the vast majority of producers of digital photographs adhere to these social conventions, and this file format offers them a way to demonstrate how their work represents, reflects, or adds to our understanding of the world. For consumers of digital images, this format elevates both the meaning inherent in the image(s), and their ability to have confidence in their relationship to the real world.

10.1 Staged development

The framework will be developed in 4 stages.

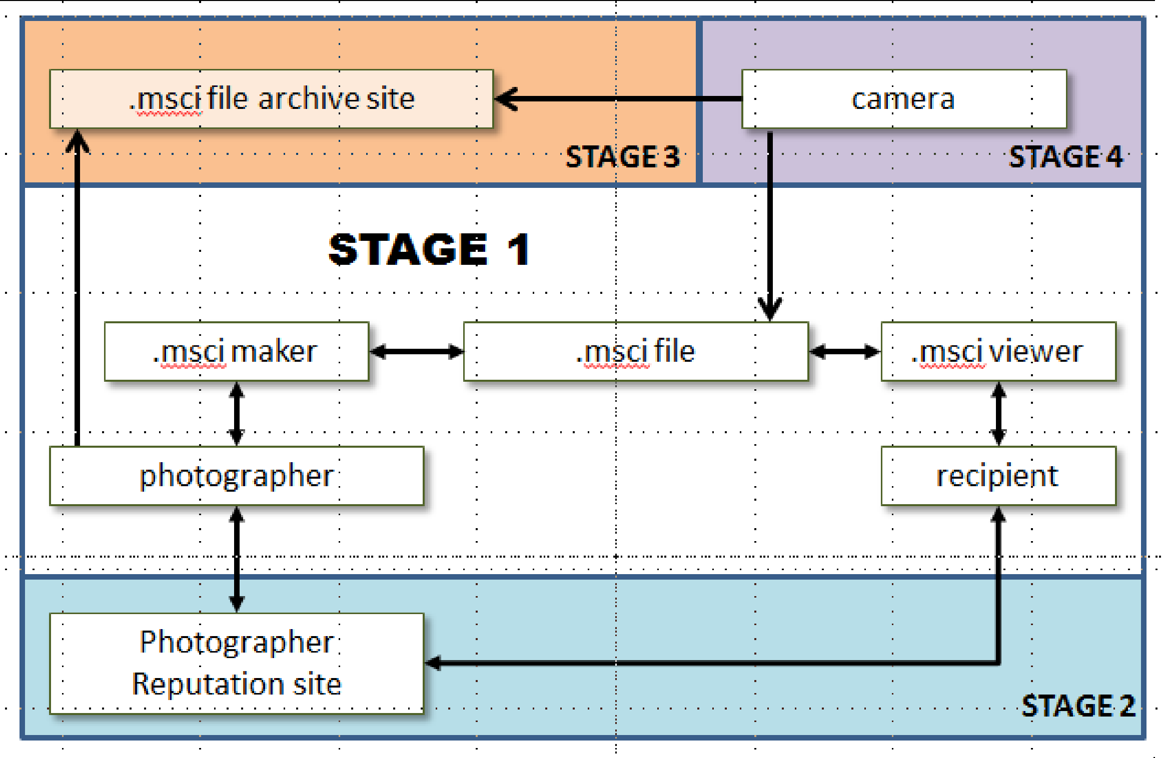


Figure 10-1: Four stages of image authentication framework

10.1.1 Stage 1 - .msci solution

The .msci solution is at the heart of the framework and will be developed first. This version of the framework design specification focusses on this stage.

10.1.2 Stage 2 – Online photographers' reputation site

To tackle the problem of 'sourceless' images, the framework will link a photographers' reputation site to the .msci system such that viewers of .msci files will be able to further ascertain the likelihood of any manipulations by having reference to the photographer's online social profile.

This site will allow photographers to upload correct copies of the .msci files they have distributed, and at the same time will allow people to leave relevant comments similar to eBay.

Funding would likely be provided via offering the site as a free for basic services and annual fee for premium services basis.

10.1.3 Stage 3 – Online .msci archive site

In addition to uploading .msci files to their photographer's reputation site, photographers will be able to upload these files to a .msci archive. This format of archive is already successful (Getty Images, stock photo sites) although no site currently focusses on image authenticity.

Funding would likely be provided via offering the site as a free for basic services and annual fee for premium services basis. Alternatively, it may be possible to partner with a large camera manufacturer such as Canon to achieve levels 3 and 4 simultaneously.

10.1.4 Stage 4 - Camera-generated .msci file

A pre-processing step performed in camera is envisaged in which the camera manufacturer creates an initial .msci file containing at least the photograph and the photograph metadata. mapped to the Proof and Provenance layers respectively, with the Proof layer replicated in the Presentation layer as the starting point for photoprocessing by the distributor:

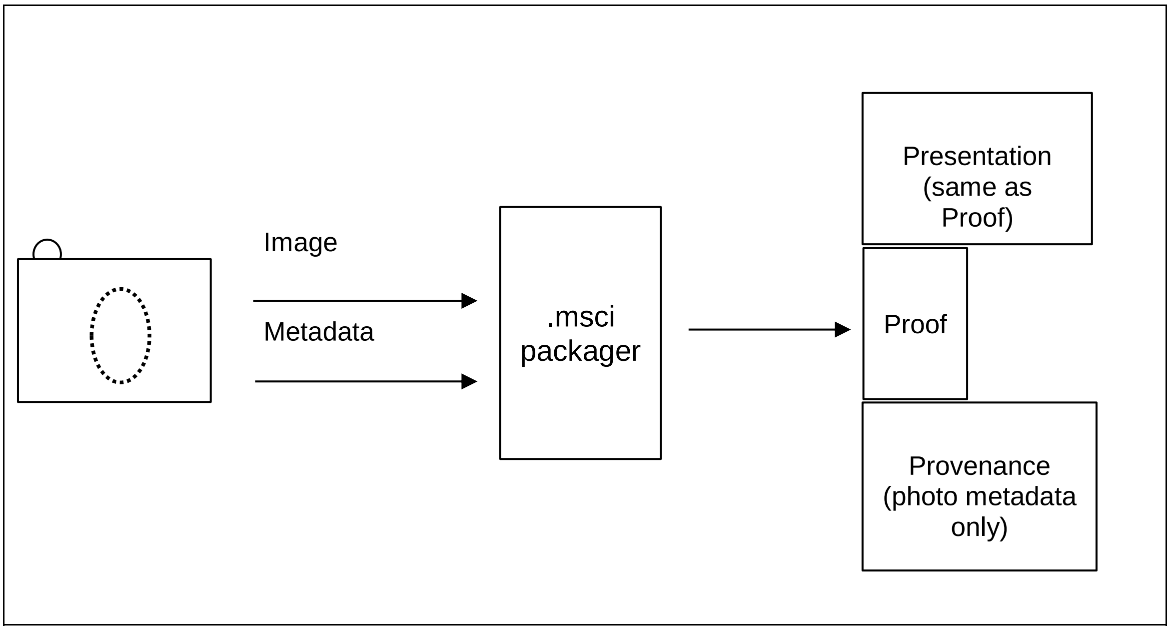


Figure 10-2: Proposed future interaction between camera and .msci

In the case of a photograph, the pre-processing element of the .msci workflow requires a camera manufacturer and therefore this element will be included in a later phase to this project. As part of this Phase I proof of concept, this step will be substituted by a 'dummy' pre-processed camera output.

However, it is important to note that this file format can accommodate a range of 'original images,' e.g. in the case of the image being a scientific graph or similar, the Proof will be a certified version of the graph by its author or in the case of the image being a scanned artwork, the Proof will be the scan, certified by the artist.

10.2 Backward and forward compatibility

The file format will accommodate backward compatibility (pre-existing images can be ingested into the Presentation and Provenance) and forward compatibility (mobility and self-containment useful as the world increasingly goes 'mobile.')

11 Architectural strategies

The development for the system will commence with the .msci suite and extend outward into the additional elements. User testing will take place locally with beta testing limited to an interested user group.

While funding is important at each stage, funding for stage 3 in which .msci file archiving will be offered will be particularly critical to enable adequate start-up and ramp-up infrastructure .

Stage 4 will require partnership with a camera manufacturer, and it may be that stages 3 and 4 can be combined.

11.1 Future plans for extending/enhancing the software

11.1.1 Self-opening .msci file

The present architecture calls for users of the system to download the .MView in order to view .msci files. In future it would be preferable that the .msci file is either able to be opened by other common programs or to be self-opening similar to the way that a Microsoft Powerpoint presentation can be saved as a self-opening file.

11.1.2 Client-side notation

Optional for consideration in a later phase: The consumer may, without starting a new file, be able to enter additional information in the provenance, including text, icons, hyperlinks and logos.[7]

7 It may be a good idea to enable audio, so that a consumer can record a comment that is added to the image package.

12 System architecture

Ultimately the framework will comprise an interconnected system of online and offline systems centering on the construction, distribution, storage and use of images packaged with meaningful comparison images and useful metadata.

12.1.1 Three tier architecture of .msci suite

The architecture of the .msci suite is three tier.

<table>
<tr><th>Tier</th><th>.msci maker (mPack)</th><th>.msci file</th><th>.msci viewer</th></tr>
<tr><td>Tier 1 (High)–
Application Level</td><td>Example(s)
Display difference map for chosen original/presentation pair of images.</td><td rowspan="3">Self-contained file created by .msci maker and opened for viewing by .msci viewer</td><td>Example(s)
Display image upon picking from array of thumbnails.</td></tr>
<tr><td>Tier 2 (Medium)
–
File Level</td><td>Example(s)
Choose images to include by picking from list
Take metadata and 3 image files, produce 1 file</td><td>Example(s)
Extract metadata from .msci file</td></tr>
<tr><td>Tier 3 – (Low) –
Libraries, bits & bytes</td><td>Example(s)
ImageMagick, XML,
Write to file</td><td>Example(s)
ImageMagick,
XML,
read from file</td></tr>
</table>

Figure 12-1: Three tier architecture

12.1.2 Programming environment and libraries

The programming environment for the .msci system is Python 1.5.2. C++ was trialled to develop the system but it was determined to be unnecessarily heavy in light of the fact that libraries such as Imagemagick exist that can be employed in providing the needed functionality to the lighter weight Python.

Platforms: It is expected that the platforms for the system will be Windows 7 and above in the first instance and OSX and above shortly thereafter.

The current version of ImageMagick for 32-bit computers is 6.8.4-10; a 64 bit version is available. ImageMagick, can be freely used in both open and proprietary products.8 ImageMagick is required for CImg to work with important ubiquitous image formats: jpg, gif, png and tiff. Natively, CImg can work with raw and bmp formats.

Potential libraries for further investigation:.

- Display thumbnails
- Ask user if they want to ingest any additional photographs
- Fullscreen display
- Create repository for notes

Global Function Requirements (in progress)

- Ability to retain connection to separate presentation image files and reload upon opening
- Ability to display thumbnails of images in a 'film strip'
- A key requirement is that the original image is not able to be changed or edited in any way. This requirement will be addressed once the general structure of the MPack, MFile and MView elements have been constructed.

[8] "ImageMagick is free software delivered as a ready-to-run binary distribution or as source code that you may freely use, copy, modify, and distribute in both open and proprietary applications. It is distributed under the Apache 2.0 **license**, approved by the **OSI** and recommended for use by the **OSSCC**." http://www.imagemagick.org/script/index.php accessed 12/3/2013

13 Detailed system design of .msci system

The .msci system is the core of the image authentication framework and is accomplished in stage 1 of the development as shown in Figure 13-1.

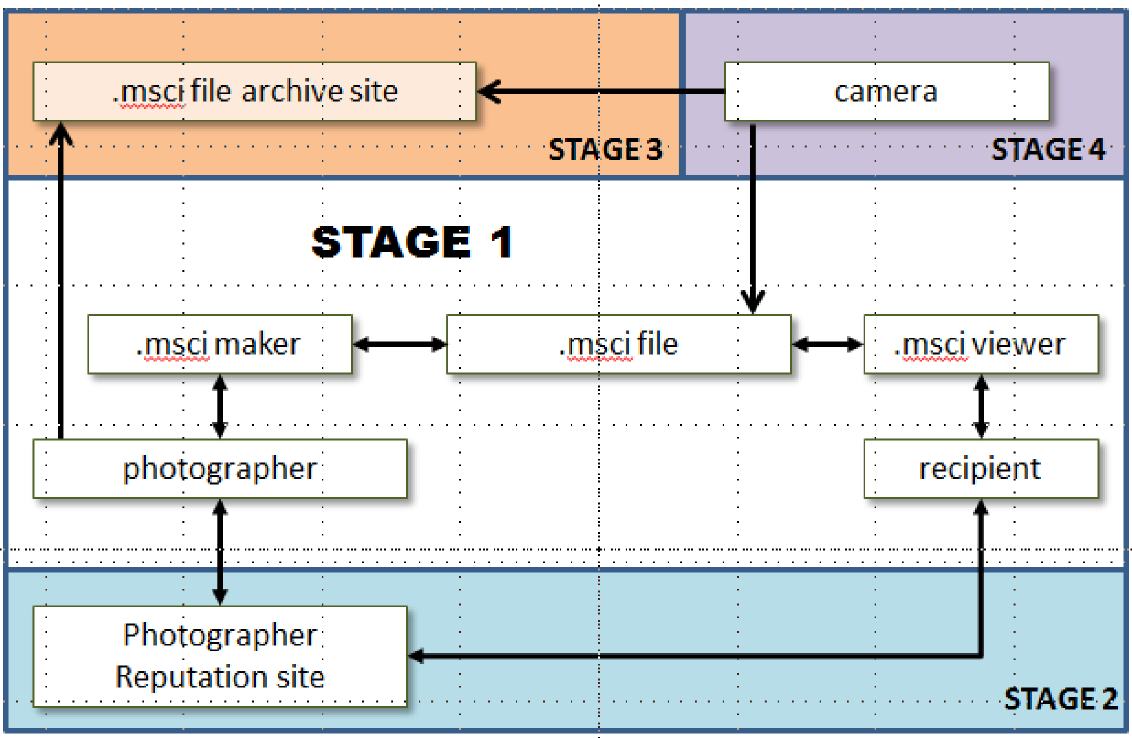


Figure 13-1: Image authentication framework .msci system components

13.1 Overview

This specification details a new file format that provides an authentication and presentation framework for original and manipulated photographs and their associated information. The name of the format is tentatively called .msci, an acronym for Mobile Self-Contained Image.

The format of the .msci file allows photographs to be distributed in such a way that the consumer viewing the image is able to easily see:

IV. The Presentation image: the image the distributor has derived from the original photograph incorporating various photo processing ranging from red-eye reduction to montage.

V. The Proof image: the original photograph(s).

VI. The Provenance: automatically captured associated information including the photograph metadata such as aperture, shutter speed, white balance, etc. and additional information provided by the distributor, such as organisation logos, standards logos, narratives, names of subjects, GPS coordinates, etc.

Thus, a .msci file can be graphically represented as trio of layers: Presentation, Proof and Provenance. In theory there may be between 0 and ∞ of each layer; in practice there would normally be at least 1 each of the Provenance and Proof layers, up to a size that is feasible for transmission between the distributor and the consumer.

Using this format an image can be transmitted as a self-contained package of information:

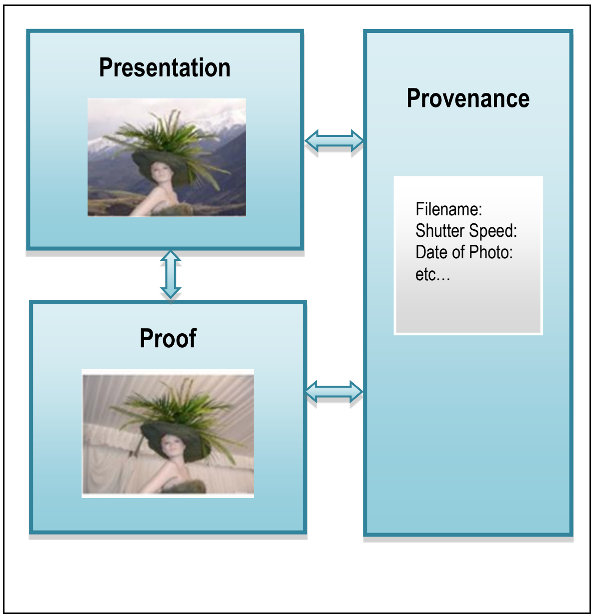


Figure 13-2: .msci file Presentation, Proof and Provenance

Presentation(s)

- Can easily be compared to proof layers
- Photo artists can show their sources
- Photo manipulations normally become evident

Proof(s)

- No Proof layer highlights 'presence of absence' and photo cannot be authenticated
- Proof(s) can be a copy, authentication string or similar, e.g. accession # of collecting institution

Provenance(s)

- Can contain easily accessible metadata
- Can hold range of contextual information
- Available for both Proof and Presentation images

The format of the file is a wrapper or container file, which allows for a diverse range of data to be included in a single file. Wrapper files also allow workflow to occur, which could be used for authentication mechanisms produced in camera and/or connected to an external locus on the interweb.

13.1.1 MPack (.msci creator)

MPack interface must be easy to use, intuitive and graphical.

MPack assembles the photo files, accepts 'ripped' information in respect of proof images, and typed information in respect of both proof and presentation images, and creates a standalone viewable file from which the original photograph cannot be removed.

Characteristics:

- Must be able to create a new MFile that includes at least one of proof and/or presentation.
- Must be able to extract metadata information from ingested photographs and display this data in a text window (tabbed in the case of multiple photographs).
- Must lock original photograph upon ingestion (to prevent purposeful or accidental modification)
- Must allow export of photographs (preferably either separately or together)

13.1.2 MFile (.msci file format)

.MFile must be secure and self-contained. It will be a wrapper/container file from which no files are allowed to be separated before during or after view.

In the event that the viewer wishes to add to the file (in the process becoming the distributor), then a new file will be created based on a copy of original file. If a viewer wishes to export images from MFile, they will require either MPack, or potentially this functionality could be built into the MFile format if it is self-opening (similar to a Powerpoint standalone presentation).

The file format will contain the elements Presentation, Proof and Provenance as outlined below.

13.1.2.1 Presentation

Characteristics: The Presentation is where the photoprocessed image is accessible for viewing.

13.1.2.2 Proof

Characteristics: The Proof is where the original photograph or other image is ingested. Once ingested, these data cannot be accessed for editing.

13.1.2.3 Provenance

Characteristics: The Provenance displays a range of textural and graphical information including metadata from the original photograph (s), additional information provided by the distributor including labels, narrative, icons, logos, and identification marks/numbers. This information cannot be accessed/changed by the consumer.9

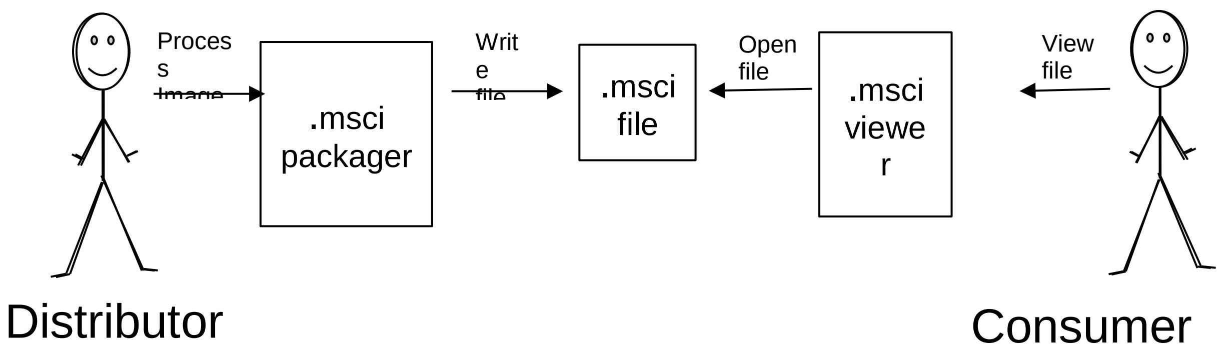
13.1.3 MView (.msci viewer)

A viewer interface for the .msci files must be non-proprietary and relatively ubiquitous OR become supported by relevant tools OR the format of the file created by the .msci packager must be self-opening (similar to a Powerpoint standalone presentation).

In the first instance the system will use a viewer application that is freely downloadable. The viewer will have minimum functionality to preserve a small footprint and to encourage users to download the Mpack application for further functionality.

13.2 Component interaction

At the top level, the system assumes a distributor, who may be an individual or an institution, and a consumer, who may also be an individual or organisation. These actors engage with the system as follows:



9 For consideration in Phase II: Provenance could include a reputation-protection model in which photographers assert authenticity, or their credibility rating is assessed externally and connected into the Provenance through workflow. Other 'soft' credibility mechanisms could be integrated in Provenance including crowdsourcing verification.

Figure 13-3: Distributor and Consumer engage with the .msci system

It may be that the consumer subsequently becomes a distributor, in the event that the consumer creates a new Presentation layer in the .msci file, potentially with the requirement to add additional layers of all three types.

13.2.1 Functional Specifications by Actor

13.2.1.1 Distributor

The distributor is the individual or organisation who transmits the file to others via posting to a website for general access or sending directly to another individual or organisation.

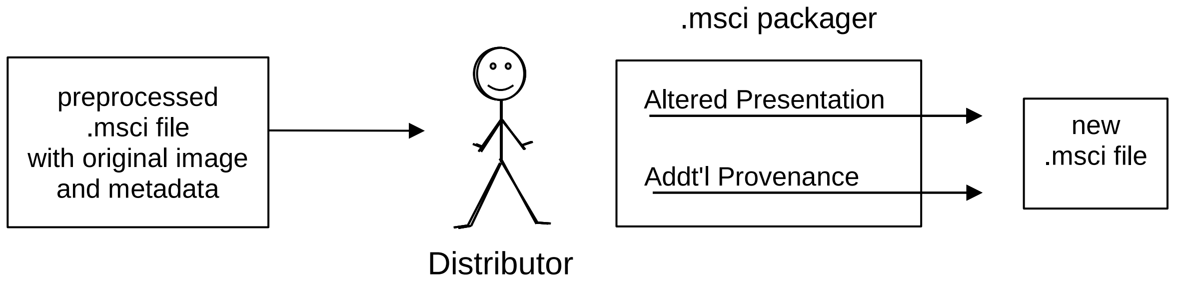


Figure 13-4: Distributor in the .msci system

Characteristics:

- Must not be able to access and/or change the original photograph appearing in the Proof or the metadata extracted from the image appearing in the Provenance.
- Must be able to add modified photographs appearing in the Presentation.
- Must be able to add information to the Presentation metadata (descriptions, purpose, copyright, etc.)

13.2.1.2 Consumer

The consumer is the individual or organisation who receives or downloads the file from the internet or email.

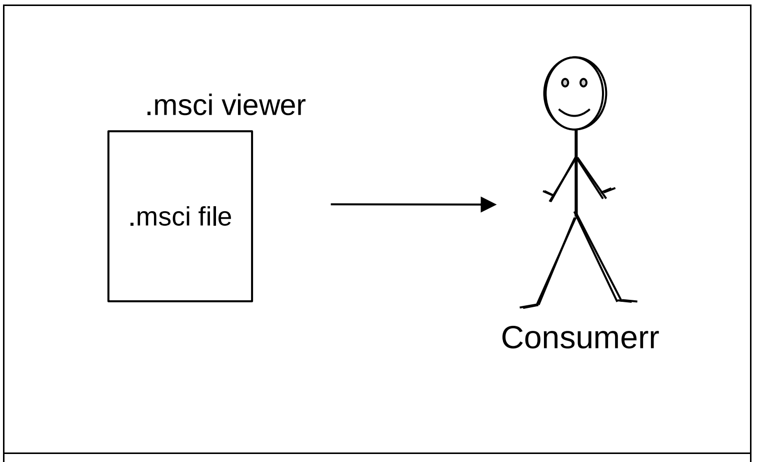


Figure 13-5: Consumer in the .msci system

Characteristics:

- Must be able to view the file without complicated steps.
- Must be able to view each image as a full screen image.
- Must not be able to access and/or change the original photograph appearing in the Proof or the metadata extracted from the image appearing in the Provenance.
- Should be able to copy the file and use as the starting point of a new MFile in the event they have installed MPack. Note that if the consumer wishes to do this, they become a distributor and all the requirements of the distributor are applicable.

13.2.2 Scenarios

Scenario 1: Robert. Robert is a freelance reporter located in Sydney. While taking photographs of the pedestrian scramble at George and Druitt streets, he snaps a photo catching a purse thief in the act. However, the scene is too large for the act to be easily seen in the photograph, so he retires to a nearby bench, opens his notebook and MPack, creates an MFile based on the original image, creates a modified image that is heavily cropped, zoomed and enhanced, attaches it to his MFile as the presentation image, jots a few notes in the text window providing more details about the event, and sends the MFile as an email attachment to his contact at the Sydney Telegraph to see if they are interested.

Scenario 2: Jeanne. Jeanne is a genealogist who works from home as a consultant investigating family trees for her clients. Today she is processing images of old worn gravestones, a rich source of familial data, from a field trip she took to the Gulgong cemetery looking for more information on the Ryan family who settled in the area in the 1860s. The gravestones are too worn to read easily or at all, and need to have significant modifications to bring the text to a level where it can be easily seen. For each relevant image, she creates an MFile containing the original gravestone photograph and the modified (presentation) photograph, then puts the MFiles in her resources directory for further research and/or transmission to her client.

13.3 .msci system content and semantic types

13.3.1 .msci system content types

The .msci file format and associated maker and viewer will manage content types including raster images, vector images and text.

13.3.1.1 Raster Images

Because the focus of this software is ‘real world’ photographs, only raster images will be supported within the presentation, proof and difference map images. Image formats are limited by the underlying libraries used by .msci such as ImageMagick.

JPG/JPEG: Required.

Logic: jpg/jpeg is the most popular format for exchanging photographic images but is lossy.

Name	**Description**	**Supported by ImageMagick?**
JPEG	24 bit RGB colour PNG without alpha (transparency) channel	Yes, requires **jpegsrc.v8c.tar.gz**

PNG: Required

Logic: png is increasingly popular because it provides efficient compression but is lossless and it also provides transparency. PNG (Portable Network Graphics ISO/IEC

15948:2003) is a raster graphics file format. There are three versions of the PNG specification: PNG: PNG-1.0, PNG-1.1 and PNG 1.2 [10]

Supported versions

Name	**Description**	**Supported by ImageMagick?**[11]
PNG	Portable Network Graphics	Yes, requires **libpng-1.2.5** or later
PNG24	24 bit RGB colour PNG without alpha (transparency) channel	Yes
PNG8	Palette-based (colour mapped) PNG images with 1, 2, 4, or 8 bit pixels. PNG8 images can be up to 256 colours and/or grayscale.	Yes
PNG32	32 bit RGBA colour PNG (with alpha channel)	Yes

TIFF: Required

TIFF (Tagged Image File Format) is a raster graphics file format. Logic: TIFF is a popular lossless format for high resolution images often available as an output format type in digital cameras. It is superior to the BMP format because it offers tags with additional information (color space, resolution, print size, etc.) and can have 16-bit colour depth.

Supported versions

Name	**Description**	**Supported by ImageMagick?**[12]
TIFF	Tagged Image File Format	Yes, requires tiff-

[10] Portable Network Graphics (PNG) Specification (Second Edition). W3C www.w3.org/TR/PNG

[11] ImageMagick format support www.imagemagick.org/script/formats.php

[12] ImageMagick format support www.imagemagick.org/script/formats.php

Name	Description	Supported by ImageMagick?[12]
		v3.6.1.tar.gz or later.

BMP

Logic: While not as powerful as TIFF, BMP is a popular lossless format with a long history in digital imaging which means there are many BMP files that could be usefully presented within an authentication framework.

Name	Description	Supported by ImageMagick?[13]
BMP	By default the BMP (bitmap) format is version 4[14]	Yes

Required Orientations for all raster image types: landscape and portrait

13.3.1.2 Vector Images

Vector images will be limited to logos and structural images.

13.3.1.3 Text / Character strings

A wide range of text and character strings will be supported including text for image metadata, user input text, labels, numeric data such as file data, filenames, etc. These will be specified during low level design specification.

13.3.1.4 Collections

Dictionaries/Lists to store metadata auto-extracted from image files. How is EXIF data read from image files? How is it handled? Where is it stored for later use?

[13] ImageMagick format support www.imagemagick.org/script/formats.php
[14] ImageMagick format support www.imagemagick.org/script/formats.php

13.3.1.5 Files

The .msci system will manage a range of file types including:

- 1 presentation image files of the types noted above.
- 0-n proof image(s) of the types noted above.
- digital signature file
- xml files
- container file

13.3.1.6 Dates and Times

Date/time data retrieved from within the image(s) metadata.

Date/time stamps for .msci file creation.

13.3.2 .msci system semantic types

The above content types will be handled in the .msci files with a range of information about the presentation and associated images.

13.3.2.1 Images

Images are ingested via the library tier and handled within the middle layer of the architecture.

13.3.2.1.1 Presentation image

This is the most significant image; it is the purpose of the communication and the main reference for most of the attendant communication (proof image, differences map image, descriptive text, copyright, etc.) within the .msci file. A presentation image is *mandatory* in the .msci file format.

13.3.2.1.2 Proof image

This is the comparator (both for ordinary visual survey and/or for a difference map) for the presentation image.

In the case of a presentation image that is unchanged from the proof image, this will be identical to the presentation image. A proof image is *optional* in the .msci file format. This is because it is expected that many images already exist that must be categorized as presentation images in the absence of any corroborating evidence of a proof image.

However if there is no proof image this must be stated (i.e. in place of proof image an application generated statement such as “No Proof Image available”). Multiple proof images will be accommodated.

13.3.2.1.3 Differences map image

The difference map image identifies the difference(s) between the presentation image and proof image. It is created by the .msci generator application using ImageMagick.

The difference map image is optional in the .msci format. However unlike a non-existent proof image, it is not necessary to display an application generated statement such as “No difference map image available”).

13.3.2.1.4 Application interface graphics

The application will use a range of graphical elements including background texture clipart, drawing clipart, and logos.

13.3.2.1.5 Photographer reputation badge

To connect the technological element of image authentication to the social element of image authentication the application interface will incorporate a system of photographer reputation badges or similar (this links to authentication framework in which visual representations of photographer reputation badges will be held outside the .msci file on the social site. The .msci file will hold a string identifying the badge separately). Note that effecting this is dependent on entering stage 2 of the development of the framework.

13.3.2.1.6 User graphics

Users such as collections archives, government departments or corporations may wish to use their own logos for display in distributed .msci files and it is important that this is accommodated.

13.3.2.2 Text / Character strings

13.3.2.2.1 Application interface text

This is text that appears within the user interface such as labels.

Figure 13-6: 'No Road'

13.3.2.2.2 User input

Assertion of digital image credibility issues. For example, a photographer may retouch a photo and state that s/he has done so.

Value-added information provided by the photographer to enhance the viewer's understanding of the image(s). For example in **Error! Reference source not found.** a research photographer might add the following description: "*No Road*: a name on an old brick pillar under the Melbourne below a tramway or rail bridge in the city of Melbourne."

Copyright information is also user input text / character string data; it provides details of type of copyright asserted by photographer such as Creative Commons or corporate licensing.

13.3.2.2.3 Filenames

Filenames will be displayed in the .msci viewer in a manner similar to that portrayed in the mock-up of the file format shown above.

13.3.2.3 Hyperlinks

There will be at least one hyperlink in any .msci file, which is the URL of the site relating to the social element of the image authentication framework. Further hyperlinks may arise from user input.

13.3.2.4 Digital signature

The digital signature exists to make any changes to a completed .msci file 'tamper evident.'

13.3.2.5 Integers

Integers are used in data associated with images including file size and pixel dimensions.

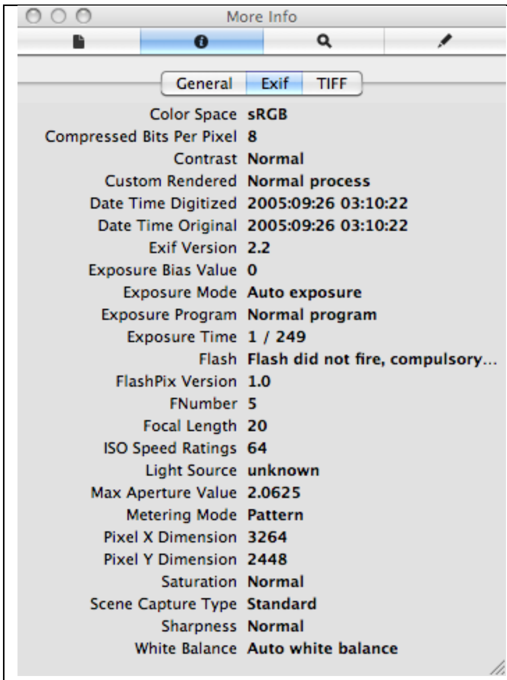


Figure 13-7: EXIF Data for 'No Road'

13.3.2.6 Collections

13.3.2.6.1 Image metadata

Cameras capture a wide range of detail about an image when the photograph is taken. These metadata are very useful in understanding the photograph but in general have a low visibility to the end user. The .msci format will make these metadata easily visible to users.

An example of the EXIF metadata associated with **Error! Reference source not found.** appears in **Error! Reference source not found.**. These data could be handled as a list.

This type of metadata is frequently overlooked by users. The .msci file will display this information in a prominent text box.

Specifically, the types of metadata found in the specified image formats such as noted in Figure 13-7 will be mapped using XML and provided with (mouseover?) explanatory help text.

13.3.2.6.2 Files

The presentation and proof images will be files of the types noted above in Content Types. The differences map image will also be an image file.

The digital signature will be a separate file. The files will be packaged within a container file similar to that in Figure 13-8.

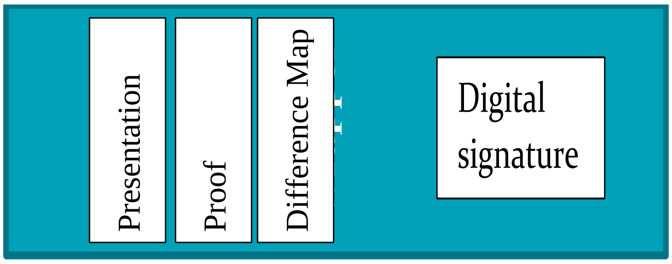


Figure 13-8: .msci wrapped file

13.3.2.7 Dates and Times

There will be date/time data retrieved from within the image(s) metadata.

There will be a .msci file creation date/time.

These date/time data will allow viewers to know when a file was created, and when the images contained therein were created.

13.3.3 MPack

MPack is an application that builds an MFile based on an initial image and its metadata. Once ingested, this original image (s) cannot be removed or altered.[15]

MPack is a Windows-based GUI environment, event-driven and navigated through the menu system and onscreen buttons.

The menu system comprises:

Menu	*Menu Item*	*Action*
File	New	Creates a new MFile, either based on an original image or 'blank.'
	Open	Opens an existing MFile
	Save	Saves current version of MFile and returns focus to the main window.
	Save as…	Creates a copy of current version of MFile, opens dialog box to browse to directory in which to save the new copy, and upon user input saves the file at that location and returns focus to the main window.
	Import original	Ingests a copy of the original image and then returns focus to the main window. As part of the processing of this menu item, the metadata is 'ripped' from the image and deposited in the textbox relating to the image.
	Import presentation	Attaches a copy of a presentation image and then returns focus to the main window. Optional 'rip.'
	Export original to new .msci	Processes ingested original file from the MFile, opens a dialog box where the image is to be exported, saves the image there as a new .msci and then returns focus to the main window.
	Create locked MFile	Locks an MFile from further editing (without password?) and creates the MFile for distribution.
	Exit	Exits the system

Menu	*Menu Item*	*Action*
Edit	Edit Presentation Provenance	Edits text in Provenance textbox specific to the image with the current focus.
	Edit General	Edits text in General Notes section.

[15] Note to myself: there will be reasons for having more than one original photo in the msci file. I need a procedural, then technical approach to creating this. So for example a gallery might issue an image as a .msci file which can then be merged with another .msci file. However in the very first instance I need to limit the development to one orginal photo. As part of Phase II I can include functionality for additional original photographs and other images.

Menu	*Menu Item*	*Action*
	Notes	

Menu	*Menu Item*	*Action*
View	View Full Screen	Opens current focus image to full screen display

Note that this menu may expand, for example it might be a good idea to offer a type of slide show for presentation images. This is also a candidate for an on-screen button.

Menu	*Menu Item*	*Action*
About	Online Help	Opens browser window and loads MPack window with help information. Currently opens to placeholder web location http://www.photos.com.au/MPack.
	About MPack	Opens window with information about the system including copyright.

Metadata / EXIF Data

Figure 13-9: Turtle skeleton, South Australia museum

Exif data (metadata) is quite extensive, perhaps beyond that which is sensible to display. For example, for the photograph shown, the EXIF data[16] depicted in Figure 13-10 contains 47 fields (repeated 4 times to cover different modes: colorspace, sampling-factor, modify and create to create a total of 188 fields).

This does not include any added fields a photographer may have included in their image or any other forms of metadata. For example, there is no information about the photograph itself, such as that it is of an ocean turtle skeleton taken at the South Australian Museum. Mpack will ingest EXIF data 'as is' in the first instance as this type

[16] Source: Author's photograph accessed for EXIF Data through http://exif-viewer.com/

of data will vary from image type to image type. Ultimately it is likely that a normalisation of metadata fields used in the .msci system will be desirable.

EXIF Field	**Value**
Colorspace	2
ColorSpace	1
ComponentsConfiguration	1, 2, 3, 0
CompressedBitsPerPixel	8/1
Compression	6
Contrast	0
CustomRendered	0
DateTime	2009:06:16 12:17:41
DateTimeDigitized	2009:06:16 12:17:41
DateTimeOriginal	2009:06:16 12:17:41
ExifImageLength	2448
ExifImageWidth	3264
ExifOffset	258
ExifVersion	48, 50, 50, 48
ExposureBiasValue	0/10
ExposureMode	0
ExposureProgram	2
ExposureTime	10/600
FileSource	3
Flash	16
FlashPixVersion	48, 49, 48, 48
FNumber	22/10
FocalLength	153/10
ImageDescription	
InteroperabilityIndex	R98
InteroperabilityOffset	2278
InteroperabilityVersion	48, 49, 48, 48
ISOSpeedRatings	64
JPEGInterchangeFormat	2486
JPEGInterchangeFormatLength	15069
LightSource	0
Make	SONY
MakerNote	Deleted due to space considerations,
MaxApertureValue	33/16
MeteringMode	5
Model	DSC-F828
Orientation	1
PrintImageMatching	80, 114, 105, 110, 116, 73, 77, 0, 48, 50, 53, 48, 0, 0, 2, 0, 2, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0
ResolutionUnit	2
Saturation	0
SceneCaptureType	0
SceneType	1
Sharpness	0
WhiteBalance	0
XResolution	72/1
YCbCrPositioning	2
YResolution	72/1

Figure 13-10: EXIF fields

List of Required Functions (in progress)

- Ability to open a user-specified image
- Ability to open and display an MFile
- Ability to create and detach an MFile
- Ability to read metadata from image files of at least types .jpg, .tiff, .png, gif
- Ability to write metadata to a text subwindow
- Ability to accept and store additional text
- Ability to ingest and 'lock' an image in the MFile

13.3.4 MFile (.msci files)

The MFile will be button driven. This is because the MFile needs to be self-contained outside of the MPack environment. A mock-up:

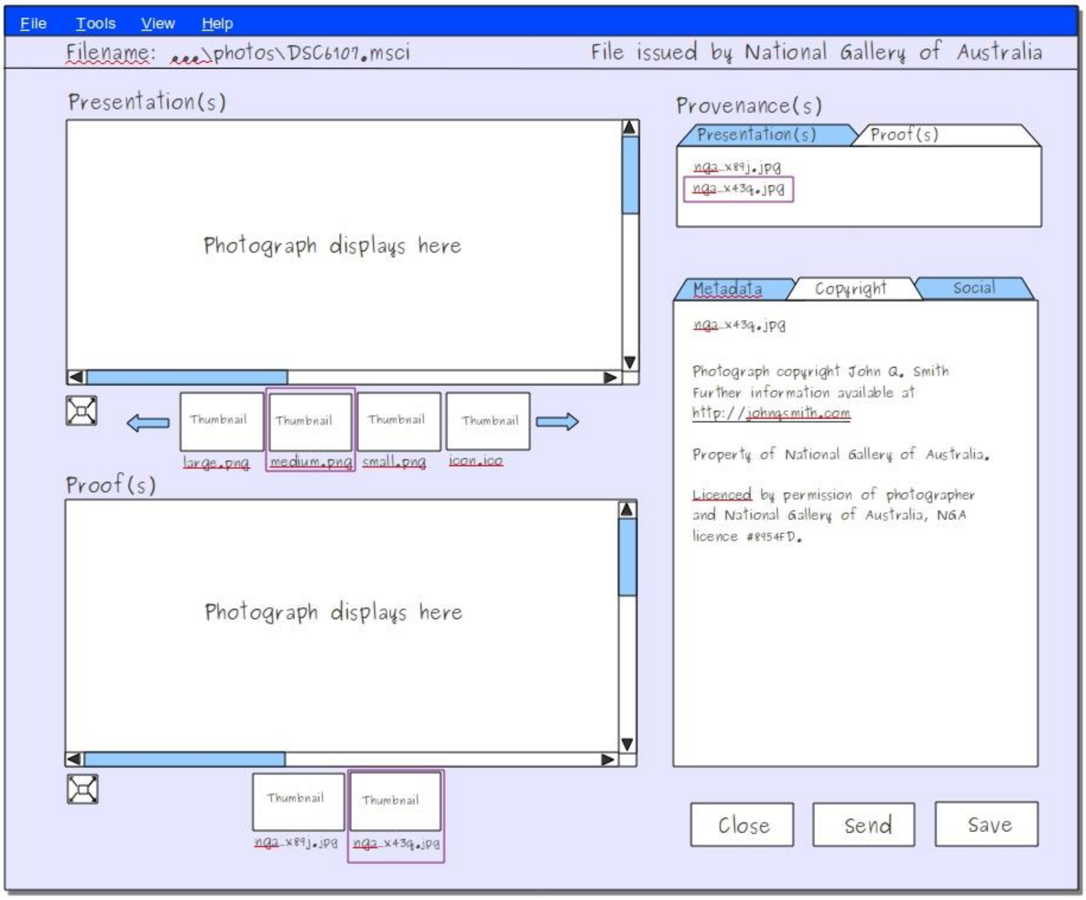
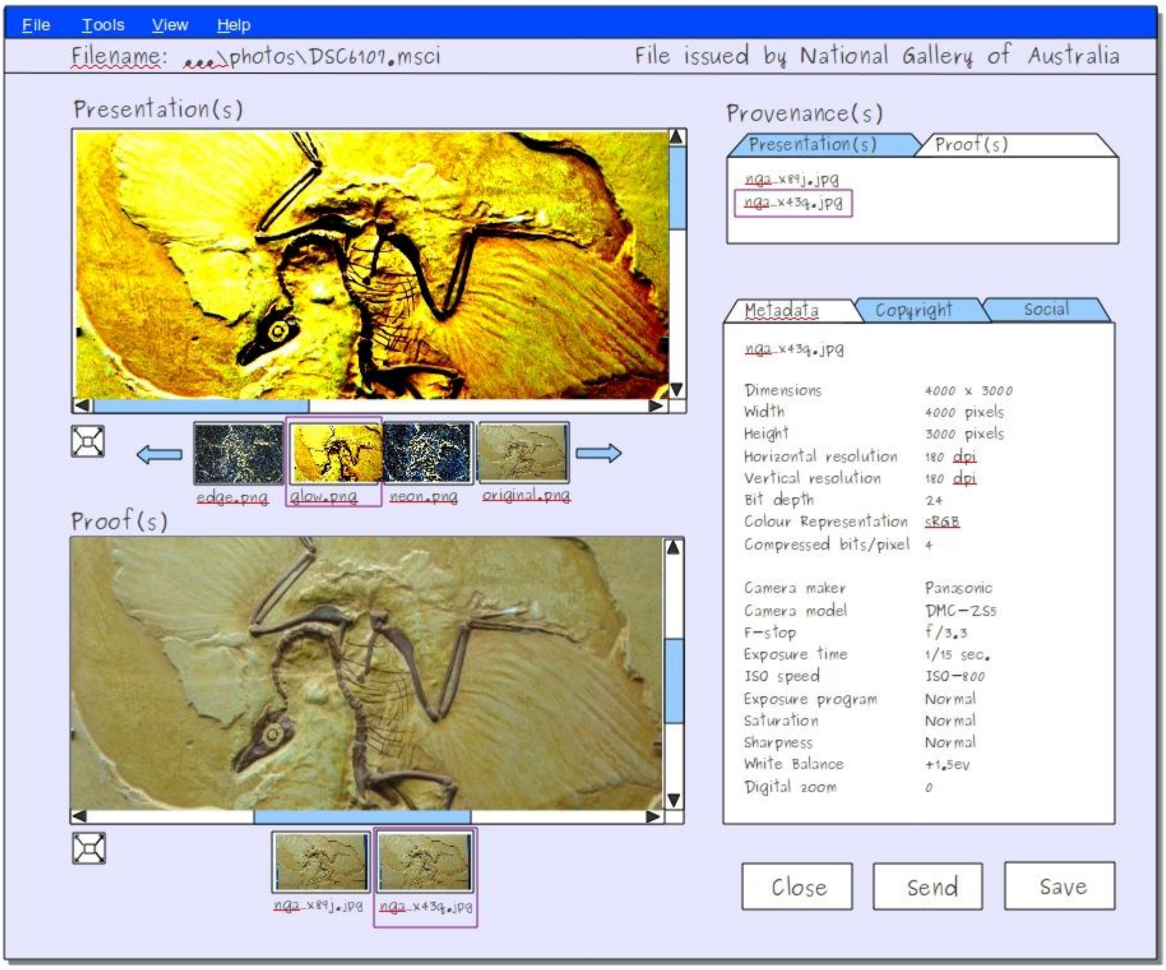


Figure 13-11: Current mockup of .msci file

In use, the Mfile might look like this:



List of Required Functions (in progress)

- Ability to open a user-specified file
- Presentation – Allow multiple images. When there is more than one presentation image, a row of thumbnail images will appear below the main presentation display area.
- Presentation – full ingestion not required, however, linked attached image files required.
- Proof – Fully ingest original image into MFile. Note limit of only one proof image.
- Provenance – Ability to rip metadata from the photograph metadata. (This can be effected through a third party tool such as Exiv2 described above or using a command line command such as the Linux *exiv2* command.)

13.3.5 MView

The MView component will be a simple display application in the first instance. Further down the track it may be possible to incorporate the ability for users to add their own notations in the file prior to further distrubiont.

14 Further work

Further aspects to be considered in this specification are:

- Distributor/consumer input interface
- Distributor/consumer output interface
- Performance requirements
- Interface requirements
- Operational requirements
- Verification requirements
- Acceptance testing requirements
- Portability requirements
- Portability to other platforms
- Quality requirements
- Maintainability requirements
- Other requirements

This page intentionally left blank.

14.1 Appendix C: Poster: Accuracy and awareness in human perceptions of veracity of manipulated and unmanipulated images

Tinkering with the Truth

sabrina.caldwell@anu.edu.au
Supervisory Panel: Tom Gedeon, Richard Jones, Martyn Jolly

Investigations into digital image manipulation and the technologies of digital image credibility

The credibility or otherwise of digital images receives little attention at present, in fact it is considered standard practice by many digital photographers and other producers of digital images to 'tinker' with digital images prior to distribution. However, as time goes on, understanding the extent to which a digital image can be believed is becoming increasingly important as 'enhanced' digital images proliferate, and purport to represent the reality of individuals and our environments, record significant human events, and are used as supportive evidence in areas as diverse as academic research and law enforcement.

Why manipulate digital images?

To exaggerate

In 2006, Reuters admitted that digital photographs presented as part of its coverage of the Israel-Lebanon conflict had been significantly altered using photo- editing software before being published. The photo above left was manipulated from the photo above right.

To vilify

In 1994, Time Magazine *digitally manipulated O.J. Simpson's mug shot to make him look more sinister. Simpson had not yet been tried and was entitled to have been presumed innocent.*

To imagine

Australian photograph "Yellow-tailed Black Cockatoos" passed on Flickr as real; photographer admits photo is same bird twice.

To lampoon

Children's book flipped to make it appear then US President George Bush is holding it upside down.

To propagandise

2008, award-winning "Qinghai-Tibet Railway opening green passageway for wild animals" photo faked by Daqing Evening News *photographer Liu Weiqing.*

Digital Image Authentication vs Digital Image Forensics

Digital Image Authentication

Authenticating copy or derivative

The image

Unmodified image

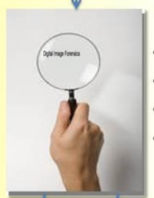
Modified image

Compare

Output authenticates or disproves image.

Digital Image Forensics

Digital image from unknown source. Examined on its merits.



- Lighting inconsistencies
- Pixel correlations
- Spliced edge detection
- Recompression artifacts

Output can only be that image is not authentic or is is *likely* to be authentic. Forensics cannot verify authenticity.

We are all affected

Courts and Law Enforcement

Digital photographs as evidence increasingly problematic.

Camera Manufacturers

Kodak, Canon, Nikon and Epson have made failed attempts to introduce authentication systems.

Scientists

In 2010, *Journal of Cell Biology* reports 25% of all images and figures submitted to them are inappropriately manipulated; 1% are fraudulent or misrepresentative.

Sociologists and Historians

"By tampering with our malleable memory, [digitally doctored photographs] may ultimately change the way we recall history," Dario Sacchi.

What's so new?

Ability to manipulate images is in the hands of the many rather than the few.

Uncritical public acceptance of photo manipulation.

In the future the period 1990's to 2020's or later may be known as the dark ages of photographs as 'silent witnesses' to truth/reality.

Conclusion

Confidence in digital images as photographs documenting human existence is eroding. Nothing less than a **robust digital image authentication process** will re-create in digital images the confidence we once placed in conventional photographs.

Australian National University

College of Engineering and Computer Science

14.2 Appendix D: Poster: Tinkering with the Truth: Investigations into digital photograph manipulation and technologies of digital image credibility

Accuracy and awareness in human perceptions of veracity of manipulated and unmanipulated images

Sabrina B. Caldwell, Tamás D. Gedeon, Leana Copeland, Richard L. Jones

Introduction

The modern world is awash in images. 350 million photographs are uploaded to Facebook alone every day.[1]

At the same time, image manipulation is achievable by anyone with image editing software.

Inappropriate image manipulation is already noted in news reporting,[2] medicine,[3] and scientific journals.[4]

We know that humans focus on salient features of an image to discern its meaning,[5] and some characteristics, especially luminance[6] attract our eye gaze.

Figure 1. ***24 Hours of Flickr*** (2013) Erik Kessels, Foam, Amsterdam[7]

But can we tell if the images we view are manipulated or not?

Methods

- 80 participants, 2 cohorts of 40 each
- Pre-familiarised participants with principles of image manipulation
- Participants viewed 14 images, 9 common to both cohorts, 5 differentiated
- Eye gaze tracked; participants queried about the images

Facelab 5.0.2 by Seeing Machines was used to track eye gaze with two infra-red (IR) cameras and a single IR light emitter pod. Eyeworks v3.8, also by Seeing Machines, was used for experiment delivery, recording and analysis.

Verbal responses analysed using Glaserian grounded theory.[8]

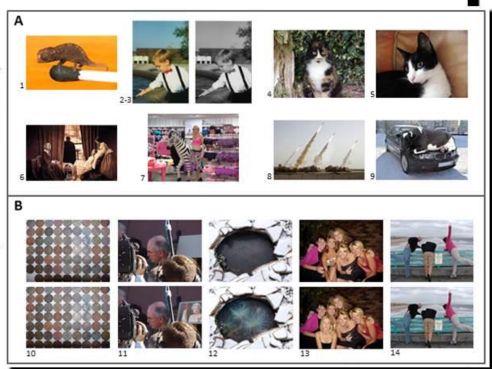


Figure 2. Common images (A) and images differentiated between cohorts (B)

Results and Discussion

Despite extensive pre-familiarisation with concepts of digital image manipulation, participants were poor to moderate at identifying whether images were manipulated or not. Participants' overall accuracy in identifying if image was manipulated or unmanipulated was 56.0% (600 out of 1071 valid image views). Accuracy in identifying what was manipulated was 27.7% (132 out of 477 valid views of manipulated images). Participants had greater success correctly identifying images as unmanipulated (61.3%) than correctly identifying images as manipulated (50.1%).

Even when participants were told directly in the case of Images 8 and 13 that a feature had been added (missile and girl added), the ability to then identify what had been added was poor: after being told an additional girl had been spliced into Image 13, accuracy increased from 11 / 40 participants to 28 / 40 participants, but 11 participants nominated other girls in the image as spliced in as well. However, we note that the average accuracy rate of participants in this experiment (56.0%) compares favourably to the authors' results in a previous experiment (Caldwell et al, 2015) in which participants with only a brief familiarisation with principles of image manipulation achieved a lower mean accuracy rate of 37.5%.[9]

Longer visual inspection of manipulated regions was associated with increased accuracy (Table 1). Comparing eye gaze tracking to question responses indicated that subjects' eye gaze fixated on regions of manipulation in images even when they did not report the image as altered.

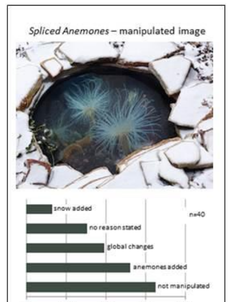
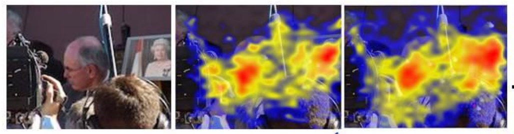


Figure 4. 67% of participants accurately identified this image as manipulated. 26% identified that the anemones were added.

Figure 3. Eye gaze 'heat maps' of Image 11 depicting former Australian PM John Howard with spliced image of Queen Elizabeth



Key: Red - most viewed, yellow - less viewed, blue - least viewed.
a) Image 11 as presented (manipulated) b) Participants stating: unmanipulated c) Participants stating: manipulated.

Table 1. Difference of participants' eye gaze viewing manipulated areas and their determinations of manipulation

Image	n	Stated: manipulated (correct)					Stated: unmanipulated (incorrect)					Sig (p<.05)	
	(total)	Mean Fixation	SD	Mean ∑ Duration (secs)	SD	n	Mean Fixation	SD	Mean ∑ Duration (secs)	SD	n	Fixation	Duration
3	63	63.4	24.43	15.39	8.29	42	56.7	25.64	15.14	8.32	21	0.329	0.908
8	67	27.0	18.07	7.66	6.08	39	28.6	23.10	8.11	6.43	28	0.765	0.772
9	67	30.2	20.72	9.60	7.56	54	29.0	17.42	8.76	6.45	13	0.837	0.689
10	20	9.3	4.35	7.75	6.68	7	3.1	2.42	2.32	2.49	13	0.009	0.002
11	37	38.5	19.09	8.34	4.77	21	25.1	12.03	5.45	3.67	16	0.014	0.045
12	35	54.5	18.82	13.02	6.65	25	32.1	15.21	6.91	5.11	10	0.001	0.008
13	34	30.1	19.32	8.24	5.97	13	30.6	18.17	8.69	5.92	21	0.941	0.836
14	36	17.0	9.40	7.54	7.93	12	12.3	6.64	4.28	2.79	24	0.143	0.192

At an individual level, the characteristics of each image (semantics and content) yielded varying outcomes aligned with the nature of the image itself. Elements of interest to participants differed for each image, and corresponding determinations of what had been changed in images they stated had been manipulated, were also image-centric (Table 2). Some themes that arose in individual image descriptions were over-arching across all images presented. In particular, the themes of uncertainty and logic presented themselves in all images. The frequency of use of these two aspects of participants' responses are detailed in Table 3 below

Image-specific themes

Table. 2. What participants stating image was manipulated perceived as being manipulated

	Image	n	Unmanipulated version	Manipulated version
Common to both cohorts	1	50	chameleon added/resized (23), background changed (19), match enlarged (4)	not presented
	2	21	colourised from black & white (10), global changes (7)	not presented
	3	45	not presented	made grayscale (32), other global changes (4)
	4	19	resized (5), cropped/zoomed (3)	not presented
	5	7	global changes (3)	not presented
	6	40	not presented	sepia/brown filter (22), man at window added (3)
	7	23	zebra added (11)	not presented
	8	40	not presented	missile(s) added (15), smoke added/changed (9)
	9	61	not presented	cow added (49), license plate blurred (9)
Differentiated	10	18	global changes (4)	global changes (6), cropped (2)
	11	29	global changes (3)	Queen added (10), global changes (3)
	12	47	global changes (13)	anemone added (12), global changes (9), snow added (3)
	13	30	global changes (6), red-eye reduction (5)	girl added (11), red-eye reduction (2)
	14	22	sign changed (6), woman added (4)	sign changed (13), woman added (4)

Overall themes:
Uncertainty and Logic

Table 3: Participants expressing overall themes of Uncertainty and Logic

Image	n	Uncertainty	Logic
1	80	24	14
2	80	12	5
3	80	13	5
4	80	14	3
5	80	10	3
6	80	11	10
7	80	5	3
8	80	20	12
9	80	22	16
10	80	21	6
11	80	20	18
12	80	18	20
13	80	16	9
14	80	27	22

Accuracy vs Confidence

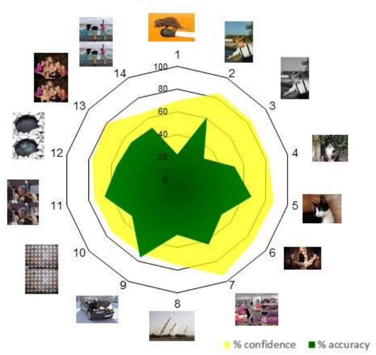


Figure 5. Accuracy < confidence < 100%

Conclusions

Being presented with a standalone image is in itself insufficient to reliably identify if the image we view has been altered. With an average success rate of 56.0%, we get it wrong 44% of the time. Even when we do successfully identify manipulated images, it is far more difficult for us to tell *what* has been manipulated (27.7%).

Increased attention (fixation, duration of eye gaze) to manipulated areas of photographs tends to be associated with greater accuracy in deciding if an image is manipulated.

Our eye gaze is a partial reflection of the features of an image that we non-consciously note, but this is not always a predictor of conscious accuracy. Further, more prominent features of an image may obscure our recognition of less obvious manipulations.

It can be inferred that we are generally poor at determining if and how the images we use in almost every walk of life have been manipulated.

It seems plausible that it is necessary for images to be accompanied by a source of additional information such as one or more of an assertion of the status of the image (manipulated or not and how), metadata, context, reference images, or verbal description.

References cited

1. Facebook, Ericsson and Qualcomm, A Focus on Efficiency. White Paper, September 16, 2013.
2. Wheeler, T. (2002). Phototruth or photofiction?: Ethics and media imagery in the digital age. Mahwah, NJ: Lawrence Erlbaum Associates
3. Prasad, H., Wanjari, S., and Rujkumar, P. (2011). Global manipulation of digital images can lead to variation in cytological diagnosis. Journal of Pathology Informatics Vol2 (20)
4. Cromey, D. (2010) Avoiding twisted pixels: ethical guidelines for the appropriate use and manipulation of scientific digital images. Science and Engineering Ethics Vol 16 No. 4, 639-67.
5. Itti L., and Koch C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Research, 40, 1489-1506.
6. Harding, G., and Bloj, M. (2010) Real and predicted influence of image manipulations on eye movements during scene recognition. Journal of Vision, 10(2):8, 1-17.
7. Kessels Kramer 24 Hours in Photos. http://www.kesselskramer.com/exhibitions/24-hrs-of-photos. Consulted 22 July 2015.
8. Glaser, B. (2007) Remodeling Grounded Theory. Historical Social Research, Supplement. No.19, 47-68
9. Caldwell, S., Gedeon, T., Jones, R., and Henschke. M. (2015). Comparing eye gaze tracking to reported perceptions of manipulated and unmanipulated digital images. Australian Journal of Intelligent Information Processing Systems. Vol 14 No 3.

Further information

For further information please contact one of the authors:

Sabrina Caldwell sabrina.caldwell@anu.edu.au
Tom Gedeon tom.gedeon@anu.edu.au

Citation: Sabrina B. Caldwell, Tamás D. Gedeon, Leana Copeland, Richard L. Jones. (2015, July). Accuracy and awareness in human perceptions of veracity of manipulated and unmanipulated images. Poster presented at the 37th Annual Conference of the Cognitive Science Society. Pasadena, California USA

Australian National University

Presented at
COGSCI 2015: THE ANNUAL MEETING OF THE COGNITIVE SCIENCE SOCIETY
July 23-26, 2015 - Pasadena,CA, USA

Research School of Computer Science
ANU College of
Engineering & Computer Science

14.3 Appendix E: The Case of the Venezuelan Poodle Moth

(Adapted from author's weblog The Photographicalist post December 2015)

We can find out practically anything at a moment's notice simply by Googling it. But the trouble with this type of instant information is that Internet content is by its nature highly fallible when it comes to accuracy. Put another way, there's a lot of fiction mixed in with the facts. And because we are usually moving quickly through the information presented to us, we can misunderstand but still think we know.

Take the case of the Venezuelan Poodle Moth for example. When I Googled this interesting sounding animal, the snap 'info window' on the right hand side of the search results (Fig 1) displayed several images and a short Wikipedia blurb.

Fig. 1: Google/Firefox Venezuelan Poodle Moth info window 5 Dec. 2015

My first impression of this moth was gleaned from this info window - the photographs, the authoritative comment from Wikipedia, the taxonomic classification. My reaction: could this moth be any cuter? I don't think so. I instantly wanted to have them in my garden, or keep one as a pet. And to think that Nature dreamed up something so adorable, that even flies! Or did she? --- Well, maybe. It is an open question, and it depends on which image you are looking at. Let's do a bit of data mining on the visual and textual information supplied in Google's info window and see if my first impressions of this moth were based on fact or fiction.

The images

There are seven images (not counting the 'People also search for' section). Two are identical, and the main photo and photo at top right are very similar. Are they all images of the Venezuelan Poodle Moth? No.

For starters, if you're looking at the gorgeous fluffy white moth with the black striated antennae in the main image, or the smaller image at top right, you aren't looking at a real Venezuelan Poodle Moth, or even any sort of a real moth.

You're looking at a beautiful moth sculpture made of wool felt. What's more, you're not even looking at a wool felt sculpture of a Venezuelan Poodle Moth, but a *Bombyx mori*, or Silk moth.

This cute little piece of art is part of a larger art sculpture (Fig 2) exhibited at the *Itami City Museum of Insects* in Hyogo Japan in 2008.[9,10,12].[17] It appears to be by a Japanese artist named Hakoiri. You can see this and more of his sculptures here .

[17] The almost intelligible Google Translate translation of his artist's statement is "Mozomozo – worms, insects Exhibition" was produced in order to participate in the exhibition that, for the first time and made insects moth of the first issue also Fumofu in wool felt, are you silkworm like that silkworm adult. And make try to, I'm more insect of wool felt facing is the Was … cute I think I think we do not, silkworm. By Chimachima flocked to thin legs, it was also representing the Fumofu feeling. Apparently so people mother's generation is Kuwabata around us until the time of your elementary school was a lot, but I'm willing to talk about the silkworm it's mon were grown until it emerged in the pupae from larvae, I brought up unfortunately It has never been. It was for sure, and vowed to mind someday."

Fig. 2: Wool felt sculpture of silk moth lifecycle by Hakoiri. Clockwise from left: caterpillar, pupae, male, eggs, female all on bed of Mulberry leaves (silk caterpillars' favourite food) [3] Photo by Tumblr photographer 'filmskiandwhatnow'

Beautiful art, but scarcely evidence of a new species of real moth, and rather misleadingly included in the information with which we have been presented.

The remaining images are real moths. Some have been taken by Dr Arthur Anker, the zoologist Wikipedia identifies as the discoverer of the Poodle Moth. Some have not. Are they Poodle Moths? Mostly no.

If you are looking at the photo of the moth in the middle top row to the right of the main picture, you are looking at a Muslin moth (Diaphora mendica) [4].

Fig. 3: Muslin Moth-photo by Flickr photographer Dr PhotoMoto [4]

This moth is associated with the Venezuelan Poodle Moth because some experts think they might be related, and also because the photographer named the moth as a 'Poodle Moth' on his Flickr site in addition to its correct name.[4]

If you are looking at the photo of the moth in the right bottom row to the right of the main picture (Fig. 4) you are looking at an unidentified moth by an unidentified photographer. It looks a bit like a portrait of an Emperor Gum moth to me, but only an entomologist can know for sure.

Fig. 4: Unidentified moth by unidentified photographer

I was unable to find any evidence of this photo having been taken by Dr Anker, despite reviewing his entire collection of Lepidoptera photos on his Flickr site. This moth photo appears to have become associated with the Venezuelan Poodle moth through sites that are erroneously including this photo as an example of the Venezuelan Poodle moth.

If you are looking at the photo of the moth in the middle bottom row to the right of the main picture, you are looking at a moth that *was* photographed by Dr Anker; he calls it simply a "Cute Moth" (Fig. 5).

Fig. 5: "Cute Moth" - Dr Arthur Anker Flickr site

This moth has also become associated with the Venezuelan Poodle moth through co-location on sites ranging from Dr Anker's Flickr photostream to quasi-scientific news sites that have been including this photo as an example of a Venezuelan Poodle Moth.[5,6] But Dr Anker does not claim this to be a Poodle Moth.

Lastly, if you are looking at the remaining two identical photos (Fig 6), you *may* be finally looking at a Venezuelan Poodle Moth.

Fig. 6: "Poodle Moth" by Dr Arthur Anker, Gran Sabana Venezuela

This photo is by Dr Anker, who states the photo date was 1 January 2009, and can be seen in various resolutions on Dr Anker's Flickr site here.

Although there is no reference original image to identify any manipulations, this image is at least cropped. The moth is obviously quite hairy; with two of its legs crossed in front of it and 5 of its 6 legs showing in the photo, it takes on an extra level of fluffy cuteness. The image caption by Dr Anker is "Poodle moth (Artace sp, perhaps A. cribaria), Venezuela."

At last we have come to the one unique photograph of the moth in question! We have to adjust our understanding of the Venezuelan Poodle moth to just that one photo. And although it isn't quite what we first expected from our Google snap info, let's face it, it is rather adorable.

And actually, does bear a marked resemblance to a poodle (for a moth).

Poodle Moth

Toy Poodle

The text and taxonomy

The text blurb accompanying this extraordinary range of images reads:

"The Venezuelan Poodle Moth is a possible new species of moth discovered in 2009 by Dr Arthur Anker of Bishkek, Kyrgyzstan in the Gran Sabana region of Venezuela. Wikipedia.

Higher Classification: Artace
Rank: Species"

This looks very authoritative, doesn't it? And frankly, like Mulder in The X-Files, I want to believe. It is evident from his 134 academic papers that Dr Anker is a respected zoologist, particularly in the field of crustaceans, especially shrimps.[11] He can therefore be assumed to be acting in good faith in presenting this moth photo as evidence of a possible new species.

But in reality, at the present time there is currently no formally identified moth by the name of the Venezuelan Poodle Moth. There is only a photo of a moth with a whimsical name. The moth is very far from scientifically described.

We have no information on its size or other physical details, its lifecycle or its specific habitat. We do not have a specimen in any collection as far as we know. Without even the most basic of information, it is difficult to ascertain from the photo that the moth is of a previously unknown species of moth from the genus Artace.

However, I can't say it will not eventually be a Venezuelan Poodle Moth, given it is as yet unidentified and could one day be a new species, which could give Dr Anker naming rights to the moth, which he **may** then give the common name Venezuelan Poodle Moth, although I think 'Anker's Poodle Moth' has a nice ring to it too.

In summary

So, in separating fiction from fact, I would say that the Google info window results would leave the casual observer convinced that this moth exists, that it is large, fuzzy and friendly, and that it comes in a variety of colours.

As a reminder, here is the Google results window I started with:

The reality of the Venezuelan Poodle Moth is that at present it doesn't exist, although a photo of a lovely unidentified moth with the title 'Poodle Moth' from Venezuela does exist in Dr Anker's Flickr photostream. Certainly it isn't the over-the-top-cute fuzzy moth that could be held in the hand (because it was a wool felt sculpture). It also doesn't come in the orange and white variants (as far as we know) that the info window evokes.

However, Google results to the contrary, this is all we know about the Venezuelan Poodle Moth:

Zoologist Dr Arthur Anker has uploaded an image to Flickr of an unusual and possibly new species of moth he reports having photographed in the Gran Sabana area of Venezuela, and he has called it a 'Poodle Moth' (illustrated below) and tentatively suggested it to be in the Artace genus of Lepidoptera.

"Poodle moth (Artace sp, perhaps A. cribaria), Venezuela."

References for "The Case of the Venezuelan Poodle Moth"

[1] Per animaldiversity.org (if this species is new and is in the Artace genus) it would be taxonomically located in the family tree as: Kingdom: Animalia – Class: Insecta – Order: Lepidoptera – Superfamily: Lasiocampoidea – Family: Lasiocampidae – Genus: Artace Source: http://animaldiversity.org/accounts/Artace/classification/ Accessed 5 December 2015.
[2] "Venezuelan Poodle Moth" http://www.snopes.com/photos/animals/poodlemoth.asp Accessed 5 December 2015.

[3] Abad-Santos, Alexander. Venezuelan Poodle Moth is the Internet's Favorite Pet this week. The Wire. 30 August 2012. http://www.thewire.com/entertainment/2012/08/venezuelan-poodle-moth-internets-favorite-real-life-pokemon-tk/56373/. Accessed 5 December 2015.

[4] Shuker, Karl. Mystery of the Venezuelan Poodle Moth – Have you seen this insect??. 22 August 2012. http://karlshuker.blogspot.com.au/2012/08/mystery-of-venezuelan-poodle-moth-have.html Accessed 5 December 2015.

[5] 'boredpanda'. 21 more strange animals you didn't know exist. http://www.boredpanda.com/unusual-animals/. Accessed 6 December 2015.

[6] Avax News. Venezuelan Poodle Moth. 13 January 2013. http://avax.news//fact/Venezuelan_Poodle_Moth.html. Accessed 6 December 2015.

[7] Abovetopsecret. The Bizzarre yet awesome Venezuelan Poodle Moth – Facts behind the hype 6 May 2013. http://www.abovetopsecret.com/forum/thread945477/pg1&mem=ratcals. Accessed 6 December 2015

[8] https://www.flickr.com/photos/artour_a/albums/72157601355579575/page6

[9]filmskiandwhatnow.tumblr.com/post/84746005790/fyeahcutebugs-hakoiri-bombyx-mori-of-wool. Accessed 7 December 2015.

[10] Itami City Museum of Insects. http://www.itakon.com/. Accessed 7 December 2015

[11] Arthur Anker. Universidade Federal do Ceara, Labomar. ufc.academia.edu/anker. Accessed 9 December 2015.

[12] http://www.geocities.jp/mekr200/hakoiri/pg118.html. Accessed 10 December 2015.

[13] Toy Poodle. FreePik. http://www.freepik.com/free-photos-vectors/poodle. Accessed 10 Dec 2015.

14.4 Appendix F: The Case of the Cold Holstein and the BMW

(Adapted from weblog post The Photographicalist September 2015)

If you Google the short phrase 'cow on car' you are sure to see this image:

Is it real? I put the question to 80 participants as part of an eye gaze experiment in which I and some colleagues used infra-red light eye tracking to see what people looked at when they saw this and other images. At the same time, we asked if they thought the image was manipulated, and if so, how.

What do you think?

Of our experiment participants, 23 were unsure. Of the remaining 57, 11 said that the photo was unmanipulated and 46 said it was manipulated. That means that the participants were divided in opinion 34 to 46, or almost half and half (42.5% to 57.5% to be exact).

Often logic (both good and bad) was employed; several participants posited that the cow had gotten up there because it was a cold day and the car had probably been recently driven as it is not covered in snow, so the hood would have been warm. One participant of Russian origin said that a cow wouldn't usually do something like that, but in Russia "we can train them to do it." Many participants said that the car would have been dented if the cow climbed up there, and one participant with a farming background said "I know how much a cow weighs, and that car couldn't stand up to it."

The farmer was right of course; the answer is that the image is manipulated, a composite of a BMW Series 3 in a snowy field or plaza near some housing, combined with an image of a perfectly normally situated Holstein milking cow (below).

Interestingly, there is a second manipulation in the image that most people missed, but that, conversely, a few people used as their main justification for deciding the image was manipulated. Can you see it?

If you can, then congratulations, you've spotted what most people didn't. The second manipulation is difficult to see via something I like to call the 'hiding effect.' The hiding effect results from a major manipulation being so eye-catching that the manipulation with a smaller profile goes unnoticed. In this case, it is the blurred out license plate. Despite its obviousness once you notice it, the cow resting on the hood distracts your attention, and indeed in our experiment the blurred license plate only received about 3% of the total attention paid to the image.

Original image of the cow lying contentedly in a green paddock.

The photo above is the original photo of the cow on the BMW. It was uploaded to a Russian photo sharing website named Kazansoft.

The BMW image must have been taken no earlier than 2003, because this E46 version of the Series 3 was manufactured between 2003 and 2005 according to my sources in the Bimmerfest BMW community (who also commented that this BMW has some modifications, for example it may have all wheel drive and someone has added some turn signal black outs and angel eyes as well.)

The Holstein image was presumably uploaded to Kazansoft some time prior to November 2013 (possibly on cdn.acidcow.com on 29 January 2013), because on 18 November 2013 the Surrey Police in England tweeted this entertaining and now relatively well known composite image of the Holstein cleverly spliced onto the hood of the BMW with a useful weather-related warning:

This image went modestly viral, and eventually became a high profile member of Internet fauxtography (Museum of Hoaxes, 2014) .

If the percentages hold true on a larger scale to our experiment results, then of the 13,706 retweeters of the post, about 5,800 of them don't know that the image is manipulated even though many might be suspicious of it.

This page intentionally left blank.

15 References

References

Agilandeeswari, L., & Ganesan, K. (2016). A bi-directional associative memory based multiple image watermarking on cover video. *Multimedia Tools and Applications, 75*(12), 7211-7256. doi: 10.1007/s11042-015-2642-1

Aileen Jacobson, S. W. (2001). FAKING IT / cropping, changing colors and adding or removing figures are tricks of the trade. but photo manipulation is a slippery slope. *Newsday*.

Al-Sanjary, O. I., & Sulong, G. (2015). Detection of video forgery: A review of literature. *Journal of Theoretical and Applied Information Technology, 74*(2), 207-220.

Anoop, M. (2015). A review on Image forgery and its detection. *International Journal of Applied Engineering Research, 10*(55), 1779-1788.

Bajac, Q. (2002). *The Invention of Photography: The First Fifty Years* (R. Taylor, Trans.). London: Thames & Hudson.

Baldwin, B. (Writer). (2016). Brook Baldwin interviews Eric Draper: CNN.

Baldwin, G. (1991). *Looking at Photographs*. Malibu, California: The J. Paul Getty Museum.

Banks, M. (2013). Chapter 9: True to Life: Authenticity and the Photographic Image. In T. Fillitz & A. J. Saris (Eds.), *Debating Authenticity: Concepts of Modernity in Anthropological Perspective* (pp. 160-174): Berghahn Books.

Blythe, P., & Fridrich, J. *Secure Digital Camera*. White paper. Department of Electrical and Computer Engineering, SUNY Binghampton.

Bossip (Producer). (2014, 29 July 2016). Carfilter. Retrieved from http://bossip.com/1079593/pure-deception-a-shocking-gallery-of-filter-sorcery/9/

Business Wire. (2007). Twenty years of the Canon EOS system leads to this moment: Canon releases the EOS-1D mark III digital SLR: Business Wire.

Buswell, G. T. (1935). *How people look at pictures: a study of the psychology of perception in art*. Chicago, Ill The University of Chicago press.

Caldwell, S. B. (2008). *Politics of Imagination: Richard Kelly Tipping and the Art and Technology of Words, Images and Objects.* (PhD), The Australian National University, Canberra, Australia.

Caldwell, S. B., Gedeon, T., Jones, R., & Copeland, L. (2015). *Imperfect understandings: a grounded theory and eye gaze investigation of human perceptions of manipulated and unmanipulated digital images*. Paper presented at the 3rd International Converence on Human-Computer Interaction, Barcelona, Spain. avestia.com/EECSS2015_Proceedings/files/papers/MHCI308.pdf

Caldwell, S. B., Gedeon, T., Jones, R., & Henschke, M. (2015). Comparing eye gaze tracking to reported perceptions of manipulated and unmanipulated digital images. *Australian Journal of Intelligent Information Procesing Systems, 14*(3).

Canon. (2016). Canon Light Awards. Retrieved 26 July 2016, 2016, from http://lightawards.canon.com.au/faq

Couch, W. (1995). Gravitational Lens in Abell 2218.

Cromey, D. W. (2010). Avoiding twisted pixels: ethical guidelines for the appropriate use and manipulation of scientific digital images. *Sci Eng Ethics, 16*(4), 639-667. doi: 10.1007/s11948-010-9201-y
Cromey, D. W. (2013). Digital images are data: and should be treated as such. *Methods Mol Biol, 931*, 1-27. doi: 10.1007/978-1-62703-056-4_1
Culbertson, A. (2015). North koriea FAKED missile launch with 'clever photo editing'. from http://search.proquest.com/docview/1682064364?accountid=8330
Cutrone, M., & Grimalt, R. (2001). The true and the false: Pixel-byte syndrome. *Pediatric Dermatology, 18*(6), 523-526. doi: 10.1046/j.1525-1470.2001.1862009.x
Cyberlink. (2016). CyberLink's New PhotoDirector 6 Brings Pro-Level Photography Techniques to Consumers and Enthusiasts.
Dellinger, H. (1997). Words Are Enough: The Troublesome Use of Photographs, Maps, and Other Images in Supreme Court Opinions. *Harvard Law Review, 110*(8), 1704-1753.
Dolly. (2009). Beat Body Bullies: All sizes, all shapes, all airbrush free. *Dolly,* 80-81.
Domo (Producer). (2016). Data never sleeps 2.0: How much data is generated every minute? Retrieved from https://web-assets.domo.com/blog/wp-content/uploads/2014/04/DataNeverSleeps_2.0_v2.jpg
Dong, J., Wang, W., Tan, T., & Shi, Y. Q. (2009). Run-Length and Edge Statistics Based Approach for Image Splicing Detection. In H.-J. Kim, S. Katzenbeisser & A. T. S. Ho (Eds.), *Digital Watermarking: 7th International Workshop, IWDW 2008, Busan, Korea, November 10-12, 2008. Selected Papers* (pp. 76-87). Berlin, Heidelberg: Springer Berlin Heidelberg.
Facebook. (2013). *A Focus on Efficiency, A White Paper*. Ericsson and Qualcomm.
Farid, H., & Bravo, M. (2010). Photo Forensics: How Reliable is the Visual System? *Journal of Vision, 10*(7), 1231-1231. doi: 10.1167/10.7.1231
Financial Times. (2002). Greater powers to manipulate images. Retrieved from http://search.proquest.com/docview/249284067?accountid=8330
Frenda, S. J., Knowles, E. D., Saletan, W., & Loftus, E. F. (2013). False memories of fabricated political events. *Journal of Experimental Social Psychology, 49*(2), 280-286. doi: 10.1016/j.jesp.2012.10.013
Garry, M., & Wade, K. A. (2005). Actually, a picture is worth less than 45 words: Narratives produce more false memories than photographs do. *Psychonomic Bulletin & Review, 12*(2), 359-366. doi: 10.3758/bf03196385
Gedeon, T. (2008). *Neural Networks for Modeling Esthetic Selection: Neural Information Processing - Proceedings of the fourteenth International Conference on Neural Information Processing (ICONIP 2007.* Paper presented at the ICONIP 2007.
Global Polio Eradication Initiative. (2010). Wild poliovirus photographed through microscope.
Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological bulletin, 134*(3), 460-476. doi: 10.1037/0033-2909.134.3.460
Gregory, R. B. (2016). Who's the most photographed american man of the 19th century? HINT: It's not lincoln., *The Washington Post*. Retrieved from http://search.proquest.com/docview/1773652686?accountid=8330
Grimes, J. (1996). On the Failure to Detect Changes in Scenes across Saccades. In K. Atkins (Ed.), *Perception*. New York: Oxford University Press.

Groesz, L. M., Levine, M. P., & Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *The International journal of eating disorders, 31*(1), 1-16. doi: 10.1002/eat.10005

Halpern, M., & Humphreys, L. (2016). Iphoneography as an emergent art world. *New Media & Society, 18*(1), 62-81. doi: 10.1177/1461444814538632

Hand, M. (2014). Persistent traces, potential memories: Smartphones and the negotiation of visual, locative, and textual data in personal life. *Convergence: The International Journal of Research into New Media Technologies*. doi: 10.1177/1354856514546094

Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *Journal of Vision, 10*(2), 8-8. doi: 10.1167/10.2.8

Harrison, B. (2002). Photographic visions and narrative inquiry. *Narrative Inquiry, 12*(1), 87-111. doi: 10.1075/ni.12.1.14har

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498-504. doi: http://dx.doi.org/10.1016/j.tics.2003.09.006

House, N. A. V., Davis, M., Takhteyev, Y., Ames, M., & Finn, M. (2004). The Social Uses of Personal Photography: Methods for Projecting Future Imaging Applications. *Working Papers, 3, 2005*.

Hsu, Y. F., & Chang, S. F. (2010). Camera Response Functions for Image Forensics: An Automatic Algorithm for Splicing Detection. *IEEE Transactions on Information Forensics and Security, 5*(4), 816-825. doi: 10.1109/TIFS.2010.2077628

Inoue, M., Mitsukura, Y., Fukumi, M., & Akamatsu, N. (2000, 2000). *Neural net based image retrieval by using color and location information.* Paper presented at the Systems, Man, and Cybernetics, 2000 IEEE International Conference on.

International Loupe Awards. (2016). International Loupe Awards competition rules.

International Pano Awards. (2016). International Pano Awards competition rules.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research, 49*(10), 1295-1306. doi: http://dx.doi.org/10.1016/j.visres.2008.09.007

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10–12), 1489-1506. doi: http://dx.doi.org/10.1016/S0042-6989(99)00163-7

Jackson, A., Lubas, R., & Schneider, I. (2013). *The Metadata Manual*: Elsevier Science.

Jacobson, J., & Werner, S. (2004). Why cast shadows are expendable: insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art. *Perception, 33*(11), 1369-1383. doi: 10.1068/p5320

Jenkins, R. V. (1975). Technology and the Market: George Eastman and the Origins of Mass Amateur Photography. *Technology and Culture, 16*(1). doi: 10.2307/3102363

Jerebko, A. K., Barabanov, N. E., Luciv, V. R., & Allinson, N. M. (2000). *Neural-net-based image matching, Applications of Artificial Neural Networks in Image Processing.* Paper presented at the SPIE3962.

Joshi, A., Caban, J., Rheingans, P., & Sparling, L. (2009). Case Study on Visualizing Hurricanes Using Illustration-Inspired Techniques. *IEEE Transactions on Visualization and Computer Graphics, 15*(5), 709-718. doi: 10.1109/TVCG.2008.105

Kawamura, S. (1998). Capturing images with digital still cameras. *IEEE Micro, 18*(6), 14-19. doi: 10.1109/40.743680
Ke, Y., Qin, F., Min, W., & Zhang, G. (2014). Exposing image forgery by detecting consistency of shadow. *TheScientificWorld, 2014*(1), 1-9. doi: 10.1155/2014/364501
Keating, C. F., Randall, D., & Kendrick, T. (1999). Presidential Physiognomies: Altered Images, Altered Perceptions. *Political Psychology, 20*(3), 593-610.
Kee, E., & Farid, H. (2011). A perceptual metric for photo retouching. *Proceedings of the National Academy of Sciences of the United States of America, 108*(50), 19907-19912.
Kee, E., O'Brien, J. F., & Farid, H. (2013). Exposing photo manipulation with inconsistent shadows. *ACM Trans. Graph., 32*(3), 1-12. doi: 10.1145/2487228.2487236
Kirchner, M., & Böhme, R. (2007). Tamper Hiding: Defeating Image Forensics. In T. Furon, F. Cayre, G. Doërr & P. Bas (Eds.), *Information Hiding: 9th International Workshop, IH 2007, Saint Malo, France, June 11-13, 2007, Revised Selected Papers* (pp. 326-341). Berlin, Heidelberg: Springer Berlin Heidelberg.
Kirk, J. (2010). Analyst finds flaws in Canon image vreification system. *PC World*.
Klein, J. (2012). Getty Images Citizenship Report: Where We Stand: Citizenship Report 2012" (pp. 5).
Kress, G. (1996). Internationalisation and Globalisation: Rethinking a Curriculum of Communication. *Comparative Education, 32*(2), 185-196.
Lang, D. (2007). Q&A: HANY FARID ON DIGITAL MANIPULATION *Photo District News* (Vol. 27(6), pp. 18-18,20).
Lebeck, R. (1958). Elvis Presley in the US Army in Friedberg, Germany.
Lefour, V. (2015). World Press Photo award withdrawn in manipulation scandal. *Business Insider*. http://www.businessinsider.com/afp-world-press-photo-award-withdrawn-in-manipulation-scandal-2015-3/?r=AU&IR=T#ixzz3UKbWyeQt
Leyden, J. (2011). *Nikon image authentication system cracked*. London: The Register.
Loftus, G. R., Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of experimental psychology. Human perception and performance, 4*(4), 565-572. doi: 10.1037/0096-1523.4.4.565
McGregor, A. (2004). *Frank Hurley: a photographer's life*. Camberwell, Vic: Viking.
Meskin, A., & Cohen, J. (2007). *Photographs as Evidence*. University of California, San Diego. San Diego, CA, USA.
Milani, S., Fontani, M., Bestagini, P., Barni, M., Piva, A., Tagliasacchi, M., & Tubaro, S. (2012). An overview on video forensics. *APSIPA Transactions on Signal and Information Processing, 1*. doi: 10.1017/ATSIP.2012.2
Mohanty, S. P. (2009). A secure digital camera architecture for integrated real-time digital rights management. *Journal of Systems Architecture, 55*(10–12), 468-480. doi: http://dx.doi.org/10.1016/j.sysarc.2009.09.005
Museum of Hoaxes. (2014). Cow on Hood of Car. from http://hoaxes.org/weblog/comments/cow_on_hood_of_car
N. J. Wade, B. W. T., D. Heller. (2003). Dodge-ing the issue: Dodge, javal, hering and the measurement of saccades in eye-movement research. *Perception, 32*(7), 793-804. doi: 10.1068/p3470
NASA. (1969). Aldrin descends the Lunar Module ladder.

National Geographic. (2016). Yourshot competition guidelines. from http://yourshot.nationalgeographic.com/photo-guidelines/
Newstex. (2011). A room filled with 24 hours of flicr photos, *Gizmodo*. Retrieved from http://search.proquest.com/docview/903664821?accountid=8330
Ng, T.-T., Chang, S.-F., Lin, C.-Y., & Sun, Q. (2006). Passive-blind Image Forensics-15 (pp. 383-412): Elsevier Inc.
NISO. (2004). Understanding metadata. *Library Hi Tech News, 21*(9).
OED. (Ed.) (2016) Oxford English Dictionary (online).
Olivier, M. (2007). George eastman's modern stone-age family: Snapshot photography and the brownie. *Technology and Culture, 48*(1), 1-1A.
Owen, A. (1994). 'Borderland Forms': Arthur Conan Doyle, Albion's Daughters, and the Politics of the Cottingley Fairies. *History Workshop*(38), 48-85.
Pollard, A. (2014). Weird use of images poses a big question, *The Canberra Times*.
Porter, G., Tales, A., & Leonards, U. (2010). What makes cast shadows hard to see? *Journal of Vision, 10*(3), 1-13. doi: 10.1167/10.3.13
Radharani, S., & Valarmathi, M. L. (2010). A study on watermarking schemes for image authentication. *International Journal of Computer Applications, 2*(4), 24-32.
Rare Historical Photos. (2016). The Falling Soldier. from http://rarehistoricalphotos.com/falling-soldier-1936/
Reaves, S., Hitchon, J. B., Park, S.-Y., & Yun, G. W. (2004). "You Can Never Be Too Thin"; or Can You?: A pilot Study on The Effects of Digital Manipulation of Fashion Models" Body Size, Leg Length and Skin Color. *Race, Gender & Class, 11*(2), 140-155.
Redi, J. A., Taktak, W., & Dugelay, J.-L. (2011). Digital image forensics: a booklet for beginners. *Multimedia Tools and Applications, 51*(1), 133-162. doi: 10.1007/s11042-010-0620-1
Richard, G. G., & Roussev, V. (2006). Next-generation digital forensics: The digital forensics community requires new tools and strategies for the rapid turnaround of large forensic targets. *Communications of the ACM, 49*(2), 76.
Robinson, H. P. (1860). On Printing Photographic Pictures from Several Negatives. *The British Journal of Photography*, 94.
Robinson, H. P. (1869). *Pictorial effect in photography being hints on composition and chiaroscuro for photographers*. London: Piper & Carter.
Royal Household at Buckingham Palace. (2016). from https://www.royal.uk
Sasse, A. (1951). Enstein sticks tongue out: United Press International.
Seeing Machines. (2016). Seeing Machines. from seeingmachines.com
Siegel, R. (1994). Famous 'photo' of loch ness monster is a fake. from NPR http://search.proquest.com/docview/190033763?accountid=8330
Skene, A. (2014, Nov 03 2014). The rise and fall of fotomat islands, *Toronto Star*. Retrieved from http://search.proquest.com/docview/1619272097?accountid=8330
Smithsonian. (2016). Smithsonian magazine photo contest rules.
Smyth, D. (2015). Giovanni Troilo finds a silver lining after World Press exclusion. *British Journal of Photography*.
Sontag, S. (1978). *On photography*. Great Britain: Allen Lane.
Steadman, I. (2013). 'Fake' World Press Photo is't fake, is lesson in need for forensic restraint. from http://www.wired.co.uk/news/archive/2013-05/16/photo-faking-controversy

Stearnes, S. (1963). UPI. Retrieved from http://www.history.com/topics/us-presidents/john-f-kennedy/pictures/john-f-kennedy/john-f-kennedy-jr-saluting-his-father-at-funeral

Steimle, S. (1998). Will Germany's Good Scientific Practice Guidelines Prevent Fraud? *Journal of the National Cancer Institute, 90*(22), 1694-1695. doi: 10.1093/jnci/90.22.1694

Sutthiwan, P., Shi, Y. Q., Su, W., & Ng, T. T. (2010, 19-23 July 2010). *Rake transform and edge statistics for image forgery detection.* Paper presented at the Multimedia and Expo (ICME), 2010 IEEE International Conference on.

Suvarna, S. K., & Ansary, M. A. (2001). Histopathology and the 'third great lie'. When is an image not a scientifically authentic image? *Histopathology, 39*(5), 441-446. doi: 10.1046/j.1365-2559.2001.01312.x

Talbot, H. F. (1839). Some Account of the Art of Photogenic Drawing, or the Process by Which Natural Objects May Be Made to Delineate Themselves without the Aid of the Artist's Pencil. [Abstract]. *Abstracts of the Papers Printed in the Philosophical Transactions of the Royal Society of London, 4*, 120-121.

Taylor, A. (2012). 75 Years Since the Hindenburg Disaster. Retrieved from http://www.theatlantic.com/photo/2012/05/75-years-since-the-hindenburg-disaster/100292/

Telegraph UK. (2013). 'Best ever' photograph of loch ness monster revealed as a fake., *Telegraph Co UK*. Retrieved from http://search.proquest.com/docview/1439454093?accountid=8330

The Telegraph UK. (2016). Digitally altered images: famous pictures that have been manipulated using Photoshop. from http://www.telegraph.co.uk/news/picturegalleries/worldnews/8982502/Digitally-altered-images-famous-pictures-that-have-been-manipulated-using-Photoshop.html?image=2

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97-136. doi: http://dx.doi.org/10.1016/0010-0285(80)90005-5

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *The Quarterly Journal of Experimental Psychology, 59*(11), 1931-1949. doi: 10.1080/17470210500416342

Velden, v. d. (2015). Forensic devices for activism: Metadata tracking and public proof. *Big data & society, 2*(2).

Waldcock, J. (2010). Applying mathematics to digital image processing using a spreadsheet. *Journal of MSOR Connections, 10*(3), 11-16.

Wang, W., Dong, J., & Tan, T. (2014). Exploring DCT Coefficient Quantization Effects for Local Tampering Detection. *IEEE Transactions on Information Forensics and Security, 9*(10), 1653-1666. doi: 10.1109/TIFS.2014.2345479

Wei, W., Dong, J., & Tan, T. (2009, 7-10 Nov. 2009). *Effective image splicing detection based on image chroma.* Paper presented at the 2009 16th IEEE International Conference on Image Processing (ICIP).

Weissmann, G. (2006). Science fraud: From patchwork mouse to patchwork data. *FASEB Journal, 20*(6), 587-590. doi: 10.1096/fj.06-0401ufm

Welling, J. (2009). *Why photography matters as art as never before: Michael Fried in Conversation with James Welling*. Aperture, (No. 195 (Summer 2009)). New York.

Wheeler, T. (2002). *Phototruth or photofiction?: Ethics and media imagery in the digital age.* (Vol. 13). Mahwah, NJ: Lawrence Erlbaum Associates.
Williams, R. (1979). Reality...What a Concept: Casablanca Records.
World Press. (2016). World Press photo competition rules.
Yarbus, A. (1967). *Eye Movements and Vision*. New York: Plenum Press.