2011

Visualising forensic data: investigation to court

Ken Fowle
Edith Cowan University

Damian Schofield
Edith Cowan University

DOI: 10.4225/75/57b2bc7240cec

Originally published in the Proceedings of the 9th Australian Digital Forensics Conference, Edith Cowan University, Perth Western Australia, 5th -7th December 2011

This Conference Proceeding is posted at Research Online.
http://ro.ecu.edu.au/adf/96
VISUALISING FORENSIC DATA: INVESTIGATION TO COURT

Ken Fowle¹, Damian Schofield¹,²
¹secau Security Research Centre, School of Computer and Security Science
Edith Cowan University, Perth, Western Australia
k.fowle@ecu.edu.au
²State University of New York
Department of Computer Science, USA
schofield@cs.oswego.edu

ABSTRACT

Visualisation is becoming increasingly important for understanding information, such as investigative data (for example: computing, medical and crime scene evidence) and analysis (for example: network capability assessment, data file reconstruction and planning scenarios). Investigative data visualisation is used to reconstruct a scene or item and is used to assist the viewer (who may well be a member of the general public with little or no understanding of the subject matter) to understand what is being presented. Analysis visualisations, on the other hand, are usually developed to review data, information and assess competing scenario hypotheses for those who usually have an understanding of the subject matter.

Visualisation represents information that has been digitally recorded (for example: pictures, video and sound), handwritten and/or spoken data, to show what may have, could have, did or is believed to have happened. That is why visualising data is an important development in the analysis and investigation realms, as visualisation explores the accuracies, inconsistencies and discrepancies of the collected data and information.

One of the primary issues of visualisation is that no matter how coherent the data, there will always be conjecture and debate as to how the information is/has-been visualised and, is it presented in an acceptable and meaningful way.

This paper presents a range of examples of where forensic data has been visualised using various techniques and technology, the paper then concludes with a discussion of potential benefits and problems.

Keywords


INTRODUCTION

At the end of the 18th Century William Playfair, a Scottish inventor, introduced the line graph, bar chart and pie chart into statistics. He demonstrated how much could be learned if one plotted data graphically and looked for suggestive patterns to provide evidence for pursuing research. However, due to the novelty of the graphical forms, Playfair had to include extensive directions for the viewer informing them how to read the data visualised from the graphs and charts he created (Tuft, 1997). Today these graphs (and many other more complex graphical representations) are a vital and everyday part of communication in science and technology, business, education and the mass media (Cleveland & McGill 1984).

Scientists and scholars have always used graphical techniques to describe, represent, and create knowledge. Traditionally, these techniques have focused on the communication of quantitative data and information (e.g., graphs and charts) although a variety of methods have also emerged to communicate more qualitative information including behavioral maps, and perspective renderings. (Ramasubramanian and McNeil, 2004).

The human visual system has the ability to interpret and comprehend pictures, video, and charts much faster than reading a description of the same material. The human brain performs some processing early in the chain of processing visual input, this process starts in the eyes. Hence, images are interpreted much faster than textual descriptions as the brain processes the visual input much earlier than textual input. This results in the human visual system’s ability to examine graphics in parallel but humans can only process text serially (Teerlink and Erbacher 2006). Marty (2008) stated “It is not just the expedited browsing capabilities that visualization has to offer, but often a visual representation—in contrast to a textual representation”
A visualisation is an image, diagram, graphic or animation representing data that is intended to give a better understanding of that data. There are many different visualisation areas, differing mostly by the domain of the visualised information. Examples include: mathematical & scientific visualisations (results from equations and formulas); product visualisation and three-dimensional design (images, photos or computer aided design software) and medical imaging (information and images from medical machines such as magnetic resonance imaging scanners).

Visualisation, in its broadest sense, is a communicative process that relies on encoded meanings that can be transferred from creators and organizers of information to users and receivers of the same information (Shannon, 1948). Edward Tuft (1997) proposes that visualisation is as much an art as a science, where the processes of arranging data and information in order to achieve representation, communication, and explanation are consistent regardless of the nature of substantive content or the technologies used to display the information.

Within the realms of forensic science, the use of new technologies in order to gather, analyse and present evidence is of the utmost importance in the modern world. Better collection and analysis of evidence from a wide range of digital media can be achieved by the use of data from the devices of perpetrators, victims and witnesses involved in incidents. The devices which may provide additional evidence include mobile phones, PDAs, tablets, digital cameras, computers and closed-circuit TV. Recent terrorist event have highlighted these new forms of evidence as mobile phone images and video are collected from members of the public who were at the scene of an incident. (Schofield 2007).

This paper presents a number of different forensic visualisation methods for two specific areas: the investigation and the presentation. Each has a common thread and that is that data visualisation is still relatively new and there is further research required to establish an accepted framework of what visual is suitable and acceptable for investigation and presentation in relation to its target audience.

**WHY VISUALISE DATA**

We have all heard the expression “a picture paints a thousand words” but this is only if the viewer has some understanding of what is being presented and why it is being presented. For example, consider the image shown below (Figure 1) and its potential ability to confuse a viewer unfamiliar with the information types being visualised.

![Figure 1](image)

The image above (Wand, n.d.) shows computer network traffic in a graphical format, the data is captured from a live network interface, visualising the flow of network data between hosts, providing (at a glance) information about network usage. To the trained person in computer network traffic analysis the image has meaning and
provides displays information but to the layperson it will require the provision of a detailed explanation. How information and data is viewed, interpreted and understood depends on what is presented, to whom it is presented and why it is being presented. Visualisations are only effective when the right kind of pictorial representation is chosen and can be manipulated to show useful information (Lowman 2010).

Images are interpreted much faster than textual descriptions as the brain processes visual input much earlier than textual input.

Many disciplines are facing an ever-growing amount of data and information that needs to be analysed, processed, and communicated. People who have to look at, browse, or understand the data need ways to display relevant information graphically to assist in understanding the data, analysing it, and remembering parts of it.

The ability of a computer to create synthetic copies of an event or issue (whether as a static image, a plan or schematic, a computer animation or a virtual reality simulation) provides the opportunity to enhance the viewer’s current understanding. These visualisations allow users to learn, question and interact within the computer-generated environment and it provides the opportunity to make mistakes, revisit and review, without necessarily putting themselves at risk.

INVESTIGATION VISUALISATION

Analysis of digital data storage is often a key area in modern crime scene investigation, so much so in fact, that the computer is sometimes now considered as a separate crime scene. The computer may hold evidence in the form of documents, e-mail records, web history and caches, login dates and times of access, and illegal files, to name but a few. The digital evidence process has become so focused around this area, that disk analysis has become known, by some authors as ‘forensic computing’ (Schofield 2007).

Today’s digital forensic investigator has “hundreds of specific and unique application software packages and hardware devices that could qualify as cyber forensic tools… hundreds of utilities available for the same task”. (Marcella and Menendez, 2007). The basic requirement for a computer forensics tool is to convert specific files into a human readable format for analysis by a forensic investigator. This analysis is often difficult and time-consuming and involves trawling through large amounts of text. Good visual interfaces and visualisations can vastly improve the time it takes to analyse data. They can help users gain an overview of data, spot patterns and anomalies and so reduce errors and tedium (Lowman et al, 2011).

In the case of a digital forensic investigation, an investigator may need to examine the network traffic on a defendant’s computer. The investigators would begin by investigating network traffic log files taken from the computer in question. Marty (2008) reports that instead of handing someone a log file that describes how an attack happened, a picture or visual representation of the log records may be used (such as the one shown in figure 1). At one glance, the picture is capable of communicating the content of this log. Most viewers can process this information in a fraction of time that it would take them to read the original log.

In the area of forensic surveying the use of visuals to reconstruct the crime scene from all the collected and recorded information, (whether it be text, photographs, sketches, survey information - anything that can record the environment to the greatest detail possible) is invaluable. The crime scene will not be available in its initial condition forever. Evidence and information needs to be recorded before crime scene officers collect and move (thereby changing the original condition of the crime scene) items of interest.

The court is usually provided with a visual of the crime scene, a drawing (or map) based on the use of traditional techniques of protractors, rulers and has been represented as a two-dimensional diagram (Figure 2) and in the past may have been crudely drawn or plotted.
Now with the use of three-dimensional laser scanning technology, which is a combination of laser scanning surveying and digital photography. This technology captures everything in true three dimensions (in the x, y and z planes) for accurate interrogation and analysis. Figure 3 depicts a three-dimensional laser scan image; the black spot (void) is where the scanner was placed to capture the crime scene. The three-dimensional model represents a quantitative, objective database of measurements, which different operators can share for subsequent analysis.
The use of animations for investigative purposes are done to assist in proving or disproving: a hypothesis, a line of enquiry, a person’s alibi or keeping the crime scene image current in the minds of the investigators.

In this example, the image in figure 4 shows a pathology reconstruction used in a murder case to investigate the nature of a stabbing incident. The autopsy reported. the cause of death was attributed to the extensive internal bleeding caused by the stab wound to the back (measuring 3.4cm in length) which pierced the heart. It was also concluded that a large amount of force would be necessary to cause the incision to the eleventh thoracic vertebra and that the bruising to the victim’s body suggested some degree of violent struggle prior to the fatal injury (Noond et al, 2002; March et al, 2004).

The image (figure 4) shows the angle of the blade as it entered the body, cutting through the vertebra. The image (figure 5) shows a hypothetical body dynamic produced to illustrate the position of the victim illustrating the position of the body so that the damage to the internal organs matches up with the angle of the knife entry. (March et al, 2004).

Unlike the environment surrounding a road traffic accident or crime scene reconstruction, where exact, surveyed measurements are usually available, pathology or medical visualisations are often based on descriptive post-mortem findings or approximate measurements. The use of generic anatomical computer models allows the recreation of dynamic events in which wounding or damage to a human body occurs. Such a reconstruction is, by its very nature, often dependent on the knowledge, expertise and opinion of medical experts. Hence, in many of these cases the advice of the expert is seen as crucial in creating a graphical representation that accurately matches the medical opinion. However, the potential inaccuracies involved mean that these reconstructions must be viewed cautiously, and the uncertainty associated with the exact position of virtual objects must be explained to the viewer (Schofield and Mason, 2010).

All of the above information has been collected, extracted and produced by qualified people and/or experts in their respective field. They understand the visuals they create and use and understand what it is being shown. These visuals are often used as explanatory tools for juries and non-experts. However, the general public are rarely presented with these visuals without extensive expert explanation, as they may not understand the raw visualisation, misconstrue the data presented or may infer a biased view from them.

**PRESENTATION VISUALISATION**

...people who watch such television programs [CSI] regularly expect better science than what they are often presented with in courts ... In other words, CSI leads viewers to expect high-tech science and something more than the intuition of the witness, so that when in court they are presented with much lower–tech science and the witness’s subjective judgment, they are likely to find it less convincing than do non-CSI viewers (Schweitzer and Saks 2007).

When using visualisations for courtrooms, Leader and Schofield (2006) report that visual displays can often act to improve the viewer’s ability to retain the evidence, maintain an interest in the proceedings, and help them to more fully understand the nature of the evidence.
Presentation of the evidence relating to a particular incident in a courtroom can bring about the need for arduous descriptions by lawyers to get across the specifics of complex spatial and temporal data. As Burns (2001) states “The presentation typically takes the form of a report, and the scientist must be prepared to explain this report in such a way that a typically science-phobic judge and jury are able to comprehend it. Presentation is everything.

As demonstrated earlier a crime scene sketch and surveys of the crime scene are performed, but when the information is to be presented in a courtroom the once crude sketch (Figure 2) (which remains as original evidence) is often transformed into a graphical image (Figure 6) that is used to describe to the court what happened, specifically the spatial location of people and objects at particular moments in time during the event or incident chronology.

Figure 22 - Courtesy G Schofield Toronto Police

As shown in figure 1, digital forensic data when collected and visualised is readable to the trained person but for the general public and untrained the images may mean very little. In the following example figure 7, represents the number of times a known computer (collected as evidence) was used to access the internet. This information is visualised using a heat-map, which is a graphical representation of data where the values taken by a variable in a two-dimensional table are represented as colours, in this case shades of grey. The information presented is easy to read and that it is clear to see that Monday–Wednesday between 05:00-09:59, and Saturday–Sunday 00:00 – 06:59 was the most common time the suspect was on the internet. The most activity was during the period between 09:00 – 09:59.

<table>
<thead>
<tr>
<th>Time</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00-00:59</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01:00-01:59</td>
<td></td>
<td>89</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02:00-02:59</td>
<td></td>
<td></td>
<td>21</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03:00-03:59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>04:00-04:59</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>138</td>
<td>39</td>
<td>25</td>
</tr>
<tr>
<td>05:00-05:59</td>
<td></td>
<td>29</td>
<td>20</td>
<td>11</td>
<td>5</td>
<td>117</td>
<td>22</td>
</tr>
<tr>
<td>06:00-06:59</td>
<td>45</td>
<td>25</td>
<td>33</td>
<td></td>
<td></td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>07:00-07:59</td>
<td>56</td>
<td>44</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>08:00-08:59</td>
<td>122</td>
<td>218</td>
<td>138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:00-09:59</td>
<td>180</td>
<td>300</td>
<td>300</td>
<td>221</td>
<td></td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>10:00-10:59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
</tr>
</tbody>
</table>

Stephenson v. Honda Motors Ltd. of America (Cal. Super. Case No. 81067) is generally accepted to be the first case to admit evidence using a computer game engine (real-time simulator). The attorney convinced a California Superior Court of the need to use the visual component of a virtual reality simulation to help a jury understand...
the nature of the terrain over which an accident victim chose to drive her Honda motorcycle. Honda argued that the terrain was obviously too treacherous for the safe operation of a motorcycle, and that, while two-dimensional photographs and videos would help provide the jury with some idea of what the terrain was like, a three-dimensional, interactive simulator was much more realistic. In allowing the evidence, the court determined that the three-dimensional view was more informative, relevant, and probative (Schofield and Mason, 2010).

Fowle et al (2001) reported that one of Western Australia’s first computer animation was when three men faced Judge Mary-Anne Yates in the Perth District Court in 17th June 1997 (R v Marotta, Bull and King) for rape. Part of the defense case was a reconstruction of an event using video and computer graphics. The video was used to document how the measurements were taken of the scene. The computer graphics were used to show that it was impossible to re-enact the allegation without certain parameters being removed. These physical constraints can be seen in a still from the computer graphic animation (Figure 8)

The advent of better technology has improved the ability to produce better quality images. The following case involved matching a possible bullet trajectory from the police marksman outside the front of the building to the victim, who was stood at a first floor window. The images in figure 9 demonstrate the level of photorealism it is possible to achieve using modern computer graphics software. The image on the left of figure 9 is a photograph of the exterior of the building where the shooting incident occurred, the image to the right shows the virtual reconstruction of the building environment (Schofield, 2007)

It is important to realise that the use of such computer-generated presentations in a courtroom is only the current manifestation of evidence illustration and visualisation in a long history of evidential graphics used in litigation (Schofield and Mason, 2010). However, computer animations and interactive virtual simulations are unparalleled in their capabilities for presenting complex evidence. The use of such enabling visualisation technology can affect the manner in which evidence is assimilated and correlated by the viewer; in many instances, it can potentially help make the evidence more relevant and easier to understand (Schofield, 2009).
DISCUSSION

"It is difficult to determine how well twelve untrained, underpaid and usually inconvenienced strangers comprehend and utilise the evidence they hear in court, especially in cases where the evidence is provided by highly trained experts such as forensic scientists.” (Wheate, 2006)

Our culture is dominated with images whose value may be simultaneously over-determined and indeterminate, whose layers of significance can only be teased apart with difficulty. Different academic disciplines (including critical theory, psychology, education, media studies, art history, semiotics etc.) help explain how audiences interpret visual imagery (Speisel, 2006). However we must always weigh the value that technology brings against the challenges that may be presented.

In some instances, it does not make sense to use technology for the sake of technology. However, as many lawyers and expert witnesses continue to push towards the dynamic presentations of video, text, documents and other forms of evidence, it seems likely that these forensic virtual models will become a more pervasive and effective alternative to the sketched, drawings and photographs traditionally used to portray demonstrative evidence in the courtroom. (Liscio, n.d)

It could be said that when visualising data, a person must have the knowledge of the data they are visualising but they must also have knowledge of how to apply the visualising techniques for their audience. Marty (2008) supports this reasoning, he reports that most people who are trying to visualise data have knowledge of the data itself and what it means, even if they do not necessarily understand the visualisation. The viewer tends to visualise only the information collected or generated by a specific solution.

Investigators face several problems when using the non-visual tools to analyse data. The main problem is information overload. The general public have difficulty in understanding large amounts of data and visualisation is particularly apt for addressing this issue. Hence the use of visualisation is often seen as a necessary tool (Lowman (2011).

The use of three-dimensional computer-models allows for the recreation of an incident illustrating the chronological sequence of events and how they occurred. However, such a reconstruction is, by its very nature, often dependent on the knowledge, expertise and opinion of the experts. These must be viewed cautiously and the uncertainty associated with items position and action within the reconstruction must be explained by the person presenting the visual to the audience.

It should be noted that in both investigation and courtroom visuals there should be some concern that the investigator/reviewer will be focused on the visual images rather than the data source. Likewise, visual evidence has the potential to be particularly misleading. It is possible that people may focus on the elements that have a high degree of visual appeal. In all these situations, new visualisation techniques and products may be used inappropriately or be used to deflect the viewer’s focus away from key evidential issues.

Around the world a number of lawyers and forensic experts are already beginning to utilise ‘slick’ visuals to replace rhetoric and depend on their audience adapting a ‘seeing is believing’ attitude to persuade juries to believe in their arguments (Galves, 2000; Girvan, 2001; Spiesel, 2005; Bailenson, 2006; Speisel and Feigenbaum, 2009). Many of these graphical displays are being created by contractors and consultants who operate externally to the organisations involved in the case (e.g. judiciary, legal team, police service or forensic laboratory). Whether one likes it or not, in the future the technology used to generate computer games is going to be increasingly used to generate advanced visual evidence presentations in a number of courtroom jurisdictions around the world. It is imperative that researchers and practitioners start to examine the implications of this technology, evaluate its potential advantages and disadvantages and assess its impact in the courtroom (Schofield, 2009).

CONCLUSION

The continuing digital revolution has had an enormous impact on the way forensic evidence is collected, analysed, interpreted and presented and has even led to the defining of new types of digital evidence (for example, digital imagery and video, hard drives and digital storage devices). Much of this digital media will end up needing to be admitted into courtrooms as evidence. In most jurisdictions around the world technology can be slow to become legally accepted. It is fair to say that, in general, legislation for the admissibility of digital media usually lags behind the technological development (Schofield and Mason, 2010).
In this paper, we have presented an overview of some of the visualisation methods currently used for investigation and analysis of forensic data and how these visualisations are presented.

This paper has highlighted thematic areas where novel digital and mobile technologies may bring improvement to the forensic process. It underlines the fact that, until recently, three-dimensional forensic reconstruction techniques have been used (along with other multimedia technologies) mainly to present forensic evidence in the courtroom. The technologies have been targeted in this area due to their success in communicating highly complex, technical spatial and temporal evidential information to the general public.

The mixing of visual metaphors and modes may be potentially disorientating to some viewers. Combining abstract human representations in photo-realistic environments may provide an unnatural experience for the viewer. Fielder (2003) has commented on the way juries may be misled by the use of visual metaphors and abstract representations in forensic animations. The photorealistic rendering of components of the virtual model, may possibly lull the viewer into a ‘seeing-is-believing’ attitude, causing a potential relaxation of their critical faculties.

Modern systems for creating visualisations have evolved to the extent that non-experts can create meaningful representations of their data. However, the process is still not easy enough, mainly because the visual effects of processing, realising and rendering data are not well-understood by the user, and the mechanisms used to create visualisations can be a largely ad hoc process. (Rogowitz and Treinish n.d.)

There is a real risk with visualisation that the way in which it can be presented as evidence or expressed as a hypothesis can result in unfairness. The solution is to then explain the true import of the evidence and giving a jury and the general public a correct understanding and explanations should be accompanied by noting the problems associated with any assessment. The qualifications and assumptions must be explained and accompanied by appropriate warnings.

In relation to courtroom visualisation, computer graphics technology advances rapidly and the public, who regularly see photo-realistic computer graphics on television, expect to see their TV experience duplicated in the workplace and specifically in the modern day training tools they use. The public expects professional visual representations illustrating complex information, polished digital media displays demonstrating the location of spatially distributed data and dynamic animated graphics showing event chronologies.

REFERENCES


Phan, D., (2007), Supporting the visualization and forensic analysis of network events, A PhD dissertation submitted to the Department of Computer Science, Stanford University, December 2007


