1. Introduction

Much research and debate has focussed on the value of CCTV and the associated criminological aspects such as whether crime is reduced, or simply displaced, by its use. The aim of this paper is not to enter into this debate, but to examine the evidential value of the new digital technologies accompanying the CCTV phenomenon. Likewise, the civil liberties and human rights issues relating to the over-observation of our brave new world have been raised and discussed widely in many fora and do not require further examination in this context.

Despite CCTV cameras being prolific in UK towns since 1990, the quality of evidence for the purposes of identification can be poor. There are approximately 4 million cameras in the UK, principally in city centres, areas where crime is prevalent or in need of high level security. Whilst technology is becoming more sophisticated, many cameras are still recording to low quality VHS cassettes which, with repeated use, can result in extremely poor quality footage.

There has been an increase in the use of digital lenses to improve the resolution of images and greater storage capacities to prevent data being erased due to constraints on resources; this, however, is not necessarily matched with increased identification abilities. The prosecuting authorities are still faced with difficulties in establishing the identity of an offender from CCTV footage. This is largely due to the difficulties associated with the human pathways of identifying known and unknown suspects which will be addressed later in this paper.

In addition to the visual quality or resolution of CCTV imagery, many professional agencies involved in the handling of evidence are now finding that they must be aware of the potential distortion of images through compression, distribution and manipulation into different storage formats. Although this is not a major concern within this paper, it must be acknowledged as a potential area for the alteration and contamination of evidence; both intentionally – in order to deface the image – or unwittingly – whilst processing the footage.

2. Identifying the Crime

The panoptic spy of local authorities and police forces can often provide damning evidence of the actus reus of criminal assault, theft or damage to property. Imagery, rather than spoken or written word, can be more compelling and persuasive of violence. Slow motion and repeated viewing may increase the perceived severity and deliberateness of actions and the abuse of the demonstrative power of visual images can be of concern in the courtroom. The images (although static, not video
feed) shown in figure 1 can convey information more accurately than a narrative describing the same situation. The quality of the CCTV evidence need not be an issue; the acts are still clearly visible in the low resolution image shown on the right.

![Image](https://via.placeholder.com/150)

Figure 1. High and low resolution images of anti-capitalist demonstrators destroying a McDonald’s sign in Barcelona. © Wikipedia (GNU Free Documentation Licence).

Should the perpetrator’s face be indiscernible, or obscured as seen in figure 1, there may be no formal evidence of identification. The prosecution would have to rely on eyewitnesses, forensic evidence or a confession to place the accused at the locus of the crime. The malleability of eyewitness memory has been demonstrated in numerous studies (e.g. Loftus, 1979; Cutler & Penrod, 1995). This research has demonstrated that eyewitness accuracy is poor when a witness has to rely on their own memory of the person. The use of CCTV footage to verify the identity of a perpetrator would seem to overcome the issue of eyewitness fallibility. A witness would no longer need to rely on their memory of the face and would only need to match the image on the footage to their remembered image.

Instead of overcoming the issue of eyewitness inaccuracy, the use of CCTV footage has presented further inaccuracy issues. Whilst the actions and general conduct of an individual may be easily established, the quality of a recording may be sufficiently poor to prevent subsequent identification of the facial features which would in turn lead to recognition of the offender by police officers, eyewitnesses or the general public. Without the vital link between offender and suspect, the video evidence is redundant. This is the prime reason that CCTV quality should be addressed.

### 3. Identifying the Person

The timeline of events and the movement of people can be tracked using one or more CCTV recordings to establish a reliable evidential trail of events, opportunities or, indeed, alibis. Here, quality is a more serious issue as the offender must be recognisable from one scene to another; and identifiable with the accused.

A great deal of psychological research has demonstrated that identification procedures for both known and previously unknown suspects should be very different. A visual comparison of high quality imagery may suggest that a more reliable decision regarding identity will be made. However, studies indicate that when the assailant is unknown, identification is poor even when the image is of high quality, regardless of format. As such, research has demonstrated that recognising or matching unfamiliar faces even in optimal conditions is an extremely error prone process.

#### 3.1 Unknown Suspects

Research conducted by Kemp, Towell and Pike (1997) revealed that matching a photograph on a credit card to a live target was extremely error prone. Only half of supermarket cashiers in this study
performed accurately, either accepting or correctly rejecting the cards when the photograph was a high quality image. When the photograph resembled someone of similar appearance, only thirty-six percent of cashiers correctly rejected the card.

Similarly, in an experiment by Bruce, Henderson, Greenwood, Hancock, Burton & Miller (1999), participants were presented with a high quality video still of an unfamiliar male together with either a target absent or target present array. An example is shown in figure 2. Images were either matched or unmatched for view and expression and participants were required to either ‘pick out’ the target from the array, or correctly reject the array.

![Figure 2. An example of an array (similar to an identification parade) but with the target face present alongside. Image courtesy of Bruce et al, 1999.](image)

The results revealed high error rates (30%) even when the images were matched for view and expression (as above). When the viewing angle differed between target and array, performance decreased further. Henderson, Bruce and Burton (2001) extended this work to show that even when only two images were presented with a video still of the target and participants were required to identify the target, accuracy was still low. These experiments were performed under ideal conditions for making correct identifications. The conditions present in CCTV imagery are comparatively inferior, suggesting even lower success rates for matching or recognition tasks.

Irrespective of whether time constraints are in place for making visual comparisons, it appears that matching unfamiliar faces is a difficult task. With poor quality imagery, identification rates would be drastically reduced if not cut completely. Although the importance of high quality CCTV is clear, greater significance can placed on the person attempting to identify the perpetrator from the imagery.

3.2 Known or Familiar Suspects

In contrast, identification of familiar faces from CCTV footage has been shown to be extremely good, even when the footage is of extremely poor quality (Burton, Wilson, Cowan & Bruce, 1999). This investigation used low quality CCTV footage, obtained from the security camera that was located at the entrance to the psychology department at Glasgow University. In one experiment participants were presented with CCTV clips that depicted someone with whom they were familiar, walking from the door towards the camera. Each participant was asked to recognise the person from a series of clips (gait obscured, body obscured, head obscured and unedited). The overall results indicated that performance was good as 73% of poor quality targets were correctly identified.
addition, it was found that while participants did use information about gait, body shape and clothes to identify the person, participants primarily used facial information to make an identification decision. When the face was completely obscured, recognition rates were drastically reduced. This suggests that the presence of a face, irrespective of the resolution quality, is significant to familiar face recognition.

Bruce, Henderson, Newman and Burton (2001) extended this work by asking groups of familiar and unfamiliar participants to match target images (high quality video clips or stills) to high quality photographs. While viewing the target footage, participants were presented with a high quality photograph and were required to state whether the two images were the same person. The results revealed a large effect of familiarity. That is, when the participants were familiar with one or both of the images (video image and/or photograph), they were able to either correctly match the images or correctly reject them with a high degree of accuracy (90%).

Furthermore, experimental investigations by Davies & Thassen (2000) both support and extend these findings. The authors examined both recognition and recall of targets from both colour and monochrome CCTV footage. This investigation supports previous research by reporting very low identification rates (15% from memory and 30% with a still-frame image). Overall, colour footage did not improve identification rates. While it did increase the number of clothing descriptors given, fewer details were offered regarding hair information and importantly, the use of colour footage increased the number of false alarms in the identification task.

To summarise, this evidence indicates that familiar and unfamiliar faces are processed differently. Furthermore this effect is evident despite the use of different types and quality of media. Indeed, it is evident that even when we are presented with high quality images (photographs/video) we are extremely poor at accurately assessing the identity of a previously unknown person. In contrast, if we are presented with images that depict a previously known person, then we are extremely accurate at identifying the face, even when the quality of the CCTV footage is extremely poor.

4. Models of Recognition

Psychological theory provides some insight into the difficulties involved in face processing. Recognising familiar faces is a fairly robust process. We can easily recognise people that we know in different contexts and views. However, for previously unfamiliar faces, recognition can be easily disrupted by changes in viewpoint, lighting and image quality. It therefore appears that recognition of unfamiliar faces is dependent on matching or recognising superficial details of the image – picture recognition (Hancock, Bruce & Burton, 2000)

Bruce and Young (1986) described a model of the stages involved in face familiar face recognition, which highlights the difficulties involved in unfamiliar face recognition. This is partly shown in figure 3. In order for a face to become familiar, it will have been seen on multiple occasions and in different contexts. Therefore a robust representation of the face will have been formed, with a greater number of memory traces. This can be termed a structural description containing memory traces that differ in viewpoint, expression and variance in lighting conditions. This permits familiar faces to be recognised in different contexts.
In contrast, as an unfamiliar face will have been seen only once (either in a ‘live’ crime or on CCTV footage), the memorial representation will be very weak and it will contain very few traces. This brief exposure is insufficient to build up a robust structural representation. Therefore subsequent recognition of the person may be specific to the single seen image. This means that it would be much harder for a witness to recognise the face when presented with a later image of the person that differed in any way (e.g. viewpoint, expression, lighting or image quality).

5. Remedies

High quality CCTV evidence and matching tasks with persons familiar with the target are the ideal standards to ensure correct identifications and guard against miscarriages of justice. Unfortunately, these ideals are not always met. Whilst improvements can be made to the technical aspects of CCTV cameras by increasing the resolution, identification by persons familiar to the perpetrator are not always possible.

The Attorney-General’s Reference 2 of 2002 (EWCA Crim 2373) indicates the types of circumstances under which identification can be made on the basis of a photographic image from the scene of the crime. These are:

“(i) where the photographic image is sufficiently clear, the jury can compare it with the defendant sitting in the dock (Dodson & Williams);

(ii) where a witness knows the defendant sufficiently well to recognise him as the offender depicted in the photographic image, he can give evidence of this (Fowden & White, Kajalave v Noble, Grimer, Caldwell & Dixon and Blenkinsop); and this may be so even if the photographic image is no longer available for the jury (Taylor v The Chief Constable of Chester);

(iii) where a witness who does not know the defendant spends substantial time viewing and analysing photographic images from the scene, thereby acquiring special knowledge which the jury does not have, he can give evidence of identification based on a comparison between those images and a reasonably contemporary photograph of the defendant, provided that the images and the photograph are available to the jury (Clare & Peach);

(iv) a suitably qualified expert with facial mapping skills can give opinion evidence of identification based on a comparison between images from the scene, (whether expertly enhanced or not and a reasonably contemporary photograph of the defendant, provided the images and the photograph are available for the jury (Stockwell 97 Cr App R 260, Clarke [1995] 2 Cr App R 425 and Hookway [1999] Crim LR 750).”
The first instance requires suitably clear high quality imagery for a jury to identify the perpetrator with the accused. However, a great deal of research (e.g. Bruce & Young, 1986; Kemp et al, 1997; Bruce et al, 1999; Henderson et al, 2001; Hancock, Bruce & Burton, 2000) has demonstrated that unfamiliar face matching is a highly error prone process and is based on matching superficial details of the image (picture-recognition) rather than actual face recognition. This research therefore suggests that this approach is flawed, despite the presence of high quality CCTV evidence. The clarity of the imagery may well be a misleading factor that distracts from the inherent fallibility of unfamiliar face processing.

The second possibility of identification from the Attorney-General’s reference involves a familiar witness. The psychological models of recognition support the use of witnesses who have prior knowledge of the suspect. This then raises the question of when a face becomes familiar. Close overt familiarity with the subject need not necessarily be required: simply previous exposure to the person from a variety of angles or number of occasions. The Bruce & Young (1986) model clearly describes the stages of familiar face recognition. A structural description of a face is formed, the description is matched to a store of familiar faces (have I seen this face before?). If the face has been seen before, semantic information is accessed (e.g. this person works in my local shop; they’re married with two children etc). Then the final stage involves accessing the person’s name. Therefore, a person can be identified without actually accessing the name. Reporting semantically identifiable information such as address, occupation, or family composition etc. may be sufficient detail to indicate prior exposure and familiarity. From the cases referred to, this extends beyond family and friends to include members of the community, police officers and anyone who may have come into contact with the subject.

When the imagery is of poor quality, a stronger memorial trace would be required for an emphatic positive identification. Therefore it may be postulated that a stronger sense familiarity may be required for low resolution or grainy imagery.

The authors would note that the jury’s repeated exposure to the accused during a trial has not been researched. Whether this would be sufficient to develop a structural memorial trace is unknown. Such a trace may not be a particularly strong structural description as the environment or context of the face will differ in the courtroom in comparison to the crime scene. The current published research involves familiarity established prior to the entire identification task rather than part of the experimental set-up which would mimic the jury-suspect familiarity scenario.

The third option can be seen as acquisition of familiarity by repeated viewing. This may not necessarily hold true in relation to the models of familiarity described above. Faces must be learnt in order to be recognised – exposure to different angles, expressions and situations is a requirement for the production of a structural description. Repeated viewing of the same footage cannot be substituted for the memorial traces obtained from prior exposure to the subject. Should the imagery be of poor quality, repeated viewing may identify items or objects that are difficult to perceive and may not have been identifiable on a single viewing. This would not apply to face familiarity and recognition.

It is also worth noting that specialists such as police and CCTV technicians are no better at identifying either known or unknown targets than untrained participants. Whilst training in identification may increase confidence in the accuracy of face matching, the accuracy of identifications does not increase. This was demonstrated in a subsequent CCTV experiment by Burton et al (1999). In this investigation there were three groups of participants – students who were familiar with the people in the footage (1) and students (2) and experienced police officers (3) who were unfamiliar with the people depicted in the footage. Participants were shown each clip twice. After they had viewed all clips they were then presented with a series of photographs and were asked to rate whether they had seen the person in the CCTV clip. The results indicated that the students who were familiar with the targets performed well. There was no difference in performance between
the unfamiliar students and police officers, although both performed significantly poorer than the familiar group.

The final viable method of comparing a crime scene image with a suspect is through the use of an expert witness. In relation to CCTV quality issues, a further problem is posed. Should the quality be sufficiently high, the judges will most probably direct the jury to make their own comparisons, thereby disallowing the expert evidence on the grounds that no further assistance is provided to the jury.

Whilst enhancements and assessment of both high and low quality imagery may be beneficial to the identification task, the superiority of familiar face recognition cannot be superceded.

6. Summary

The quality of evidence for the identification of both known and unknown suspects is a persistent problem and remains a difficult issue for the judge or jury to establish. Quality appears to be less of an issue when familiar witnesses are involved in the identification stages of a trial, although clearly the use of high quality footage would be advantageous. Research involving unfamiliar face matching indicates that this is a difficult task. Irrespective of image quality, such a procedure should be approached with great caution and care taken to ensure that poor quality footage (and possibly high quality imagery too) is supported by other forms of evidence.

A jury may be assisted by the use of expert image interpretation: a visual comparison using a technical enhancement can result in a more sound conviction than the uncertainty surrounding grainy low-resolution images. However, high quality imagery would not generally require expert opinion, which is, in the authors’ opinion, a failure of current practices and procedure.

The legal guidance, in the form of the Attorney-General’s reference, does not appear to be in accordance with psychological research, although further studies are still required to fully understand how familiar and unfamiliar faces are recognised in the context of criminal identifications from CCTV. In conclusion, much has been learned from the overlapping interests of both psychological and legal research. The benefits, in terms of criminal justice, can be reaped by utilising the results and encouraging further research in these areas.

References