

# The importance of ground truth: An open-source biometric repository

Morgan J. Tear<sup>1</sup>, Matthew B. Thompson<sup>1,2</sup>, Jason M. Tangen<sup>1,2</sup>

<sup>1</sup> National ICT Australia

<sup>2</sup> School of Psychology, The University of Queensland

Advances in forensic technologies and procedures seek to produce better and more efficient policing for safer societies. Little is understood, however, about how effectively the human forensic professional employs such technologies, or the cognitive and perceptual processes of judgment and decision making the forensic professional engages in during the course of evidence evaluation. For this, experimenters need materials that approximate the realism of crime scene evidence, while ensuring the ground truth about the source of this information. These two goals are often incompatible. We discuss the development of an open-source biometric repository to address the issue of ground truth. This repository contains a range of crime related materials such as fingerprints and palm-prints, shoe-prints, faces, handwriting, voices, and irises. Our goal is to provide a large, open-source repository of forensic information, where certainty of the source is built into the system, to help advance research on identification by humans and technology.

## INTRODUCTION

Recent international commitments to facilitating responses to elevated threats to individuals, society, and associated infrastructures have led to changes in many Homeland Security and Policing technologies, legislation, and communications standards. There has been a large effort in many countries to develop technology for surveillance infrastructures through millions of CCTV systems and biometrics for homeland security technologies (Welsh & Farrington, 2009). In the biosecurity area, sensor networks for monitoring of critical resources such as water, soil, and the satellite monitoring of crops have also been deployed (Oliveira, Trezza, Holzapfel, et al. 2009; Ni-Meister, 2008; Wesseling & Feddes, 2006).

Although these technologies are increasingly being relied upon to capture and identify important information, they cannot effectively solve these threats without reliable, efficient, and timely interpretation. Often, the final decision is left to human judgment. Errors are to be expected, given the natural tendencies of the human practitioner for distraction, lapses of attention, fatigue, rush to judgment, analysis of imperfect information, biases, etc. (Cole, 2005; Dror & Charlton, 2006). Even the most diligent forensics professional cannot avoid all of these at any given time.

Forensic experts identify evidence on the basis of various features and rules that presumably aid classification and interpretation (Cole, 1998). Many of these encoding schemes, however, are leftovers from early, and perhaps deficient, systems. For example, fingerprint classification is based on the same general ridge formations proposed by Henry in 1900 (arches, loops, and whorls) and local features such as areas, angles and distances between ridge endings, bifurcations, forks, and enclosures (Cole, 1998). Fingerprint experts often describe prints in terms of the angle and distance between the delta and core markings, as it is assumed that these are important features for identification. These patterns look very different from one another and they contain information that is extremely difficult, for novices and experts alike, to interpret or even verbalize. Similar arbitrary coding methods can be found in face or

voice recognition, ballistics, footwear impressions and tire tracks.

## The Importance of Ground Truth

Experiments are necessary in order to investigate both the judgment and decision making processes, as well as the cognitive and perceptual processes, that underlie forensic identification. For these experiments, it is necessary utilize stimuli that approximate the high degree of visual noise, similarity, and ambiguity that professionals face when deciding whether two pieces of evidence originated from the same source. An experiment might display a crime-scene latent print alongside a fully-rolled police ten-print, with the participant required to make a decision about whether the two prints come from the same source or not. Ideally, such investigations would make use of the very same evidence that is obtained from a crime scene. The problem, however, with using actual crime scene evidence as stimuli for examining identification accuracy, is the ground truth of the evidence. That is, how can we be sure that the crime scene evidence does in fact belong to a convicted suspect? If we use stimuli for which we can guarantee such matches, the stimuli could be seen artificial. But if we use 'real' crime scene evidence, then the question of the ground truth of the evidence remains pertinent.

## The Solution

Our solution to this conundrum is to work on both fronts: accessing authentic crime-scene evidence, and creating our own stimuli where ground truth is known. Our team is working with forensic professionals across Australia to gain access to the very same materials that they routinely use in their investigations and training, where the evidence has been independently confirmed by other examiners and corroborating evidence. At the same time, we are building a standardized set of forensic stimuli that varies systematically in quality, where we have multiple pieces of biometric data that converge on a single source, and where the ground truth of the source is built into the system.

## Forensic Informatics Biometric Repository (FIBR)

The term *Forensic Informatics* refers broadly to the science of processing crime related information, where the processing is conducted by humans and technology. The advantage of using technology is obvious, but its interaction with human forensic identification has been neglected. In order to facilitate and accelerate research in the area of forensic informatics, we are currently building an open-source Forensic Informatics Biometric Repository (FIBR) that consists of a large, standardized set of biometric materials where the certainty about the source of the information is ensured.

The open-source nature of the repository will likely prove to be an important feature. While we capture and provide biometric stimuli for other researchers, we also provide protocol/instructions for others to collect their own standardized materials and contribute to the repository for the benefit of others in the field.

## METHOD

Initially conceived to facilitate research in fingerprint examination, we decided to expand the scope of our repository to include biometric materials useful to other areas in forensic decision-making. First, our collection of biometric material includes fingerprints and palm-prints, shoeprints, faces, handwriting, voices, and irises. Second, we collect the biometric data in as controlled a manner as possible while using advice from forensic professionals to simulate the nature of the evidence commonly found at a crime scene. This will allow researchers to create experiments that effectively simulate the information that forensic examiners and technologies routinely process. Finally, we collect biometric information from participants over two sessions to approximate the natural variation that is commonly found in forensic evidence (e.g., change in facial hair, clothes, shoe decay, etc.).

### Participants

Participants are first-year psychology students who participate in one hour of biometric data collection for course credit. Participants read and signed informed consent release forms.

### Stimuli Design

**Fingerprints and Palms.** Fingerprints have been used for over 100 years as a means of identification and is often regarded as the 'gold standard' of forensic science (Schwinghammer, 2004). Computer algorithms designed to aid fingerprint identification continues to advance (Alam, Akhteruzzaman, & Cherri, 2004). The final decision, however, is made by a fingerprint examiner, so it is important to approximate the same materials that they encounter. Because experts routinely match crime scene latent prints to full rolled ten-print cards, we are collecting both forms for the repository.

Latent prints vary along a number of dimensions, including quality, size, amount of useful information, and surface that they are found on. In consultation with local fingerprint examiners, we chose to use five surfaces on which latent prints are commonly found. Almost all of these

surfaces are found near points of entry to a building, and include gloss-painted timber (doors, window frames), smooth metal (door handles, knife blades), glass (windows), paper, and smooth plastic. As there are no reported figures on the variability of latent quality, latent size, and amount of useful information, we devised a protocol for collecting simulated latents, which we call 'freedom latents'. We instruct our participants to interact with the latent surfaces by asking them, for example, to 'push on the gloss-painted timber to open the door' or 'safely grab the knife by the blade.' By interacting with objects in this way, participants will leave the most realistic latents, controlling for real-world variability in latent quality, size, and amount of useful information. Two example latent prints can be seen in Figure 1.

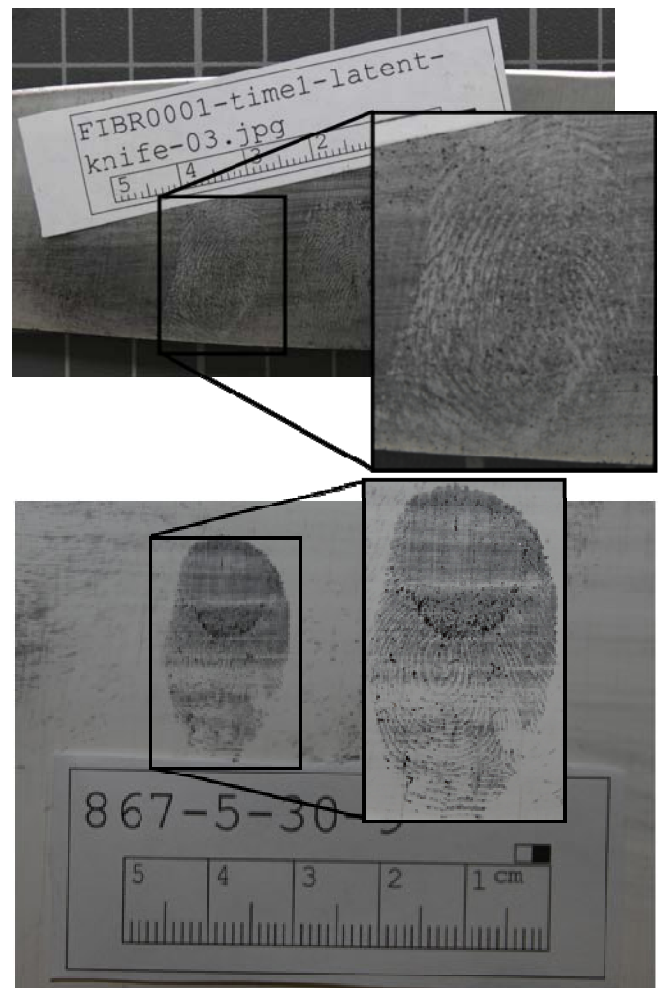


Figure 1. Example latent prints. Top: knife blade, bottom: gloss-painted timber.

We also collect standardized 10-print cards and palm-prints. The 10-print cards use ink to capture each fingerprint, rolled fully from nail-edge to nail-edge, as well as 'slap impressions' (pressing, not rolling, the fingers on the card). See Figure 2. for an example 10-print card. We separately capture fully rolled palm-prints.



Figure 2. Example 10-print form, with fully rolled prints (top two rows) and slap impressions (bottom row).

**Shoeprints.** Criminals must enter and exit a crime scene and therefore it is natural that they would leave traces of their footwear. Hilderbrand (1999) argues that when properly collected and preserved, shoeprints can provide the type, make, description, approximate size, number of suspects, path through and away from the crime scene, the involvement of evidence and the events that occurred during the crime. Consequently, the importance of footwear impression evidence should not be underestimated.

Shoeprint identification can be classified at a class and individual level. Bodziak (2000) defines a class characteristic of a shoe as an intentional or unavoidable characteristic created during the manufacturing process and is shared by one or more other shoes. He further explains that any single class characteristic will be shared by many other shoes. Individual characteristics occur when something is randomly added to or taken away from a shoe outsole that either causes or contributes to making that outsole unique. For example, cuts, scratches, tears, rocks wedged in the outsole, chewing gum, holes and air-bubbles. Bodziak (2000) argues that when sufficient, individual level characteristics are present in the impression and correspond with those on the shoe outsole, that outsole can then be positively identified as having made the impression.

Shoeprint impressions are commonly found in soft surfaces, such as mud or dust deposits. Representative latent shoeprints can be difficult to capture in a quick and clean manner. We chose to only use inked impressions to collect standardized whole and partial prints. Shoeprints vary in many dimensions. We chose, however, to vary our shoeprints along dimensions similar to fingerprints; size and amount of useful information. We vary our shoeprints by instructing participants to stand in one of three configurations, as seen in Figure 3.

Our repository also includes photographic images of each shoe that we obtain prints from. These photos are taken from multiple angles, including photos of the sole.

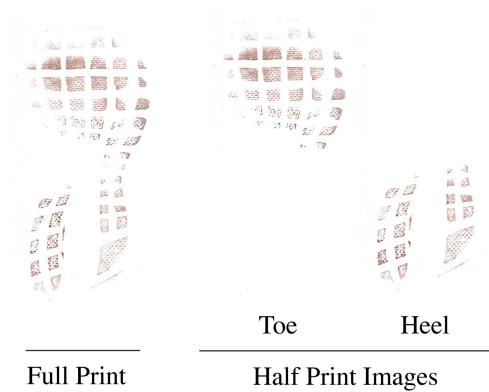


Figure 3. Partial shoeprint images.

**Faces.** Face recognition and eyewitness testimony, i.e. artists renderings, identification from mug shots and lineups, identification from review of videos, etc., have been studied extensively, and a large, high quality repository of face images and video footage is in high demand. Our repository was designed with this literature in mind and so includes images of faces from multiple angles and contexts. We collect standardized face stimuli as 'mug-shots' and crime-scene face stimuli as context face capture and mock crime footage. Our face capture protocol is divided into four stimuli types; mug-shot face capture (from five angles, see Figure 4), CCTV footage of a person entering a room, CCTV footage of a mock crime, and full-body stills.

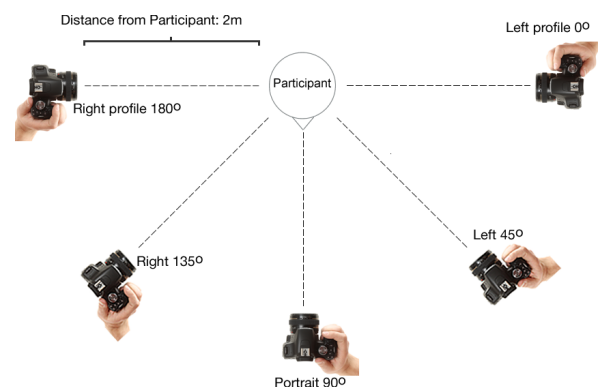


Figure 4. Angles of face capture

**Handwriting.** Forensic document examiners perform similar duties to those of fingerprint examiners, where they compare the visual features of handwriting associated with a known writer to those of an unknown writer, and then make a decision about whether the two samples belong to the same source. Handwriting samples are classified as either questioned handwriting or known/specimen handwriting. We will collect known/specimen handwriting for the repository. The nature of possible future experiments, however, means that our known/specimen handwriting samples can also be used as questioned handwriting samples when testing novice document examiners. Our handwriting protocol requires participants to write out the 'London Business Letter', which is a 'catch-all' paragraph that contains all 26 letters of the alphabet (upper and lower case) and numerals 0-9 (see Figure 5.)



"Our London business is good, but Vienna and Berlin are quiet. Mr. D. Lloyd has..."  
 Our London business is good, but Vienna and Berlin are quiet.

Figure 5. Example sentence from London Business Letter

**Voices.** Speaker-recognition is a unique area of biometrics where examiners deal with auditory noise, rather than visual noise, and so is included in our repository. The time constraint on our collection sessions, however, restricts us to only collect a short voice sample. Accordingly, participants' voices are recorded as they read out the London Business Letter.

**Iris.** Iris-recognition is a powerful method of biometric identification because of its high reliability (Daugman, 2004). False matches are incredibly rare because the iris-pattern variability among different persons is so vast (Daugman, 2004). Whereas extensive research has been conducted on automated forms of iris-recognition, there has been no work investigating the ability of humans to discriminate or match irises. We envision our iris stimuli could be used to test human ability for discriminating or matching human irises. We use a simple iris scanner to capture two standardized images of participants' irises, as well as taking 2 high quality color photos of participants' irises (see Figure 6).

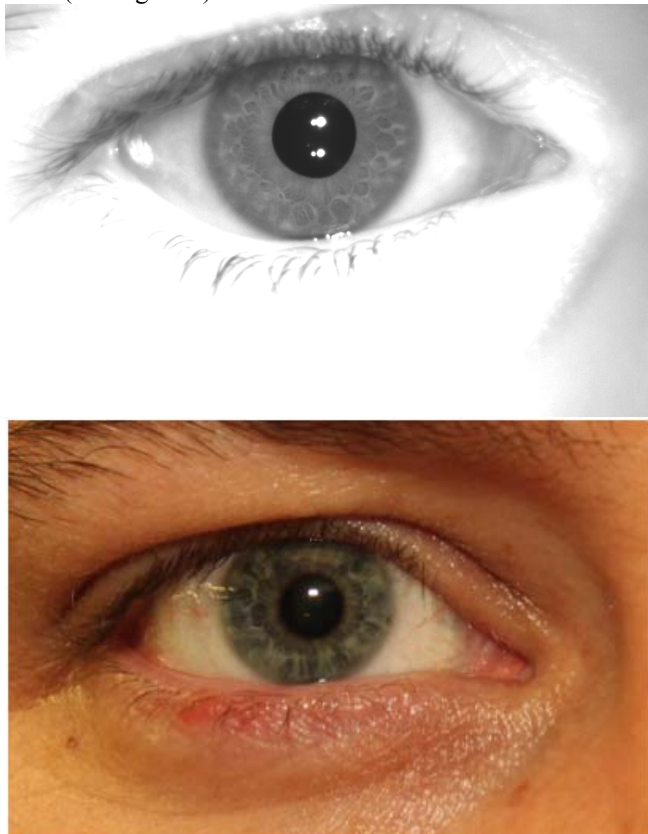


Figure 6. Top: simple infra-red iris scan, bottom: HD color photo of iris.

## CONCLUSION

In order to conduct experiments on the perceptual and cognitive processes underlying forensic identification, it is essential that researchers use materials in which the ground truth of the source is guaranteed. By creating our own repository, we have complete control over the method in which the materials are collected. On two separate

occasions (separated by several weeks), we collect digital photos of participants' face and video footage of them committing a crime, latent and full-rolled fingerprints, shoe prints, voice samples, handwriting and signature samples, and scans of their irises. In order for research in this field to progress, it is critical that all researchers in the field working on computer and human identification have free access to a large database of standardized forensic materials. The Forensic Informatics Biometric Repository (FIBR) will, therefore, be made freely available for research purposes.

## REFERENCES

- Bodziak, W.J.(2000). *Footwear Impression Evidence: Detection, Recovery and Examination* (2nd Ed.). Boca Raton: CRC Press.
- Cole, S.A. (1998). Witnessing identification: Latent fingerprinting evidence and expert knowledge. *Social Studies of Science*, 28(5-6), 687-712.
- Cole, S.A. (2005). More than zero: Accounting for error in latent fingerprint identification. *The Journal of Criminal Law and Criminology*, 95(3), 985-1078.
- Daugman, J. (2004). How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 21-30.
- De Oliveira, A.S., Trezza, R., Holzapfel, E.A., Lorite, I., & Paz, V.P.S. (2009). Irrigation water management in Latin America. *Chilean Journal of Agricultural Research*, 69, 7-16.
- Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, 56, 600-616.
- Ekman, P. (1999). Basic Emotions. In T Dalglish, & M. Power (Eds.). *Handbook of Cognition and Emotion*. Sussex, U.K.:John Wiley & Sons, Ltd.
- Hilderbrand, D.S. (1999). *Footwear, the Missed Evidence*. Wildomar: Staggs Publishing.
- Ma, L., Tan, T., Wang, Y., & Zhang, D. (2004). Efficient iris recognition by characterising key local variation. *IEEE Transactions on Image Processing*, 13(6), 739-750
- Ni-Meister, W. (2008). Recent advances on soil moisture data assimilation. *Physical Geography*, 29(1), 19-37.
- Schwinghammer, K. (2004). Fingerprint identification: How 'The Gold Standard of Evidence' could be worth its weight. *American Journal of Criminal Law*, 32(2), 265-289.
- Welsh, B.C., & Farrington, D.P. (2009). Public area CCTV and crime prevention: An updated systematic review and meta-analysis. *Justice Quarterly*, 26(4), 716-745.
- Wesseling, J.G., & Feddes, R.A. (2006). Assessing crop water productivity from field to regional scale. *Agricultural Water Management*, 86(1-2), 30-39.