Combining Forces:
Data fusion across man and machine for biometric analysis.

Sarah V Stevenage
Department of Psychology, University of Southampton
S.V.Stevenage@soton.ac.uk

Richard M Guest
School of Engineering and Digital Arts, University of Kent
R.M.Guest@kent.ac.uk

Through the HUMMINGBIRD framework outlined here, we seek to encourage a novel multidisciplinary approach to biometric analysis with the goal of enhancing both understanding and accuracy of identification.

1. Introduction

Within our modern environment, with its technological and social complexities, the capacity to recognise and identify those with whom we interact is paramount. To borrow from the recent report by the Chief Scientific Advisor to the UK Government [1], our key question is '…can I trust this person? This, in turn, depends upon two factors: are they who they say they are (authentication), and do they have the necessary permission to act as they do (authorisation)'. Biometric analysis has been pivotal in enabling answers to these questions through understanding both the measures that allow differentiation between individuals, and the means to determine a match (or mismatch) between samples. However, in many respects, standards of biometric analysis are failing to meet our needs. Two examples serve to illustrate the point: First, the misidentification of an innocent suspect is still the main cause of a false arrest or conviction within the criminal justice system [2]. This carries a myriad of consequences for the individual concerned, for the victim, and for the true perpetrator. Second, failure to spot differences between a passport holder and the image in their document remains a cause for concern at border control checkpoints [3]. With the fight against terrorism ever-prevalent, the protection of our national borders is a context in which public confidence needs to be maintained, and zero error can be tolerated.

In a world in which even the 'liveness detectors' within biometric sensors can be spoofed [4], we have to question whether our technological sophistication is protecting us from threat or exposing us to it. With this in mind, the present article provides a thought-experiment with the aim of improving biometric analysis through fusion. Current practices in biometric fusion are evaluated, however, the fusion approach is

extended beyond its existing conceptualisation providing a design framework to guide new research in identity and identification.

2. Biometric Measures in Isolation and in Combination

Biometric analysis has been applied to a number of different measures, or modalities. In this regard, it is useful to differentiate between those modalities which provide a direct record of the body (such as a fingerprint, face or footprint), and those which provide a record of the *behaviour* of that (part of the) body (such as a swipe pattern, facial expression or gait pattern). The former are often referred to as 'static' or 'physiological' biometrics whilst the latter represent 'dynamic' or 'behavioural' biometrics. In a recent report by the Parliamentary Select Committee to the UK Government [5], emphasis was placed on the value of both types of biometric when identifying an individual. To take the face as an example, whilst the static face itself has clear value when determining identity, unique identifying information may also be captured in terms of how that face moves. Consequently, when undertaking a biometric analysis of any single modality, the task is now quite considerable.

Importantly, recent approaches to automated biometric analysis have sought to improve identification performance through the application of fusion methods. Fusion has been documented through three distinct approaches all of which can be applied when evaluating whether a particular sample donated by an individual is a match (or not) to a reference sample within a database. The approaches describe (i) fusion across algorithms, (ii) fusion across presentations, and (iii) fusion across measures, and each showcase both strengths and limitations. Fusion across algorithms occurs when several algorithms are applied to analyse a particular sample. The outcome is based on a combination of the individual decisions, and the risk associated with an error determines the level of agreement sought between the individual algorithms. In contrast, fusion across presentations uses multiple presentations of the same

characteristic (for example, two index fingerprints captured separately). Samples may be fused to form a *common sample* used for comparison. By its nature, this common sample will often be more representative of the individual than the separate samples, and consequently the accuracy of the outcome will be improved. Finally, fusion across measures occurs when separate modalities are considered (for example, a fingerprint and an iris), with different tailored algorithms applied to each modality. Equally, fusion across measures may occur when different characteristics from a single modality are considered, possibly from multiple sensors. The outcome is based on a combination of the independently derived decisions to generate a normalised feature vector, enabling robust identification decision through the triangulation of information.

A number of reviews demonstrate the performance benefits (in terms of accuracy of authentication) when information is combined in this fusion approach [6, 7]. This benefit can stem from the mere triangulation of multiple pieces of information, however, the most intelligent fusion models can apply a higher weighting to some pieces of information over others such that confident decisions or effective algorithms can be valued over less confident or less effective ones. However, the argument for biometric fusion extends beyond the mere improvement of accuracy. Reliance on multiple biometrics also provides resilience against situations in which a single biometric may be unavailable (i.e., fingerprints may be worn through habit or occupation), may be temporarily unreliable (i.e., fingerprints may exhibit irregularities associated with a cut), or may be socially or culturally unacceptable to donate (i.e., religious, cultural or personal factors may result in a preference not to provide certain information). As such, a fusion approach has much to offer [8]. Our argument, however, is that examples of fusion described above should perhaps only be the start of our consideration. The remainder of this article considers the value of extending the fusion approach further to incorporate a much-overlooked factor - the strength of the human analyst.

3. The Human as a Biometric Analyst

As a highly successful social animal, the human has evolved to be able to tell con-specifics apart through the development of elegant heuristics or short cuts. These enable the human to achieve an acceptable level of

performance, with an acceptable level of speed and risk. Such short-cuts are unlikely to be captured in the code of an algorithm. Moreover, humans may also demonstrate superior performance on occasion compared to an algorithm. When considering face recognition for example, humans remain able to match a face to a reference sample even under sub-optimal conditions of altered pose, expression or lighting, which easily derail the engineered solution. Similarly, human performance may be facilitated through the completion of concurrent, complementary tasks. For example, voice recognition is improved when the voice is presented alongside its corresponding face through the interaction of the separate brain regions responsible for voice and face processing. In this sense, doing two complementary things at once can actually facilitate performance in a way that is not captured by the automated fusion across measures described above.

Given this, it is possible that the human decision maker may be able to enrich the capability of the computer-algorithm. Indeed, the criminal justice system requires that the evidence presented to court rests on the decision of a human analyst rather than being derived solely by a computer-based algorithm. Set against this, however, it is clear that the performance of a human decision maker is limited by cognitive factors and this is where the speed and impartiality of the computer algorithm offers strength. For example, the influence of cognitive load is illustrated through a deficit in human processing known as a 'dual task cost' [9]: when asked to do two unrelated things at once, the completion of the primary task impairs the ability to complete the secondary task in which case a fusion task will have defeated its purpose. Whilst this dual task cost may be overcome by splitting a complex task amongst independent human decision makers [10], the consequence is an increased requirement for man power making this undesirable as a solution. Human performance is also limited by the influence of cognitive bias [see 11]: when faced with a complex task, humans may too readily conclude that a sample and suspect are a 'match' (confirmatory bias) [12] or may be influenced by irrelevant additional information creating expectation bias, hindsight bias and contextual bias [see 13]. In all cases, human frailties can lead to errors.

4. Man-Machine Fusion

These factors suggest that the performance of a

computer-based algorithm and a human decision maker each have the capacity to be complemented by the other. The machine may benefit through heuristically derived human capacities sustaining performance in sub-optimal conditions. Similarly, the human may benefit through the application of fast, objective and unbiased decision-making free from cognitive limits. The combination of human and machine would take biometric fusion into a multidisciplinary domain, with benefits possible in terms of an enriched understanding of biometric analysis, improved resilience against unforeseen events such as cyber-attacks, as well as enhanced performance outcomes.

One example serves to illustrate the advantages gained by man-machine fusion and relates to the improvements in recognition accuracy from the voice. The strength of the machine algorithm is its capacity to process the speech signal independently of processing the semantic meaning of the speech itself. Indeed, this is something that the human listener tends to get distracted by, as shown by poorer performance in voice recognition tasks when the listener has to work hard to determine speech content. Such a situation may arise when a speaker has an unusual accent [14], is speaking in an unfamiliar language [15], or when speech is artificially disrupted through asynchrony with lip movements [16] or through temporal reversal [17]. However, one benefit that the human listener can bring to the voice recognition task is a particular advantage when processing distinctive voices. Indeed, the accuracy of automated voice matching reaches 82.8% when voices sound typical, and reaches 81.3% when voices are distinctive. In contrast, human accuracy when matching the same distinctive voices far exceeds these automated levels at 95.3%. The consequence is that automated voice recognition can be improved upon by trusting the human when the voice is distinctive [18].

5. Presentation of a Novel 'HUMMINGBIRD Framework' for HUMan-MachINe Biometric IDentification

Whilst the benefits of man-machine fusion of decision making are evident in the biometric realm, what has been missing up to now is the presentation of a formalised model to manage this fusion. We outline here the HUMMINGBIRD framework capable of organising the integration of modalities and decision makers.

This builds upon the work of Dror and Mnookin [19]

which explored how the involvement of an automated system may change the nature of the human analyst's task. For example, the involvement of AFIS in a fingerprint matching scenario may affect the level of proof required to demonstrate a match, may introduce additional bias into the human analyst's task, and may affect the probative value of the final decision depending on whether man and machine have arrived at their combined decision through repetitive or cumulative processes. Whilst these are indeed important considerations, they miss the complementary perspective of how the human analyst may shape or change the task of the machine. HUMMINGBIRD attempts to capture these.

Figure 1 shows the HUMMINGBIRD framework operating under an identification scenario. It allows for individuals to be enrolled through the donation of samples from a particular modality, however, the framework is agnostic as to which modality is used. Within a forensic context, samples will be based on available evidence, however in a biometric scenario, samples can be collected in more constrained or controlled conditions with due regard to issues of availability, acceptability, and capture. In order to successfully identify an individual, a probe sample must be compared against each of the enrolment sample clusters. The best matches become candidates for a final match/no match decision.

As in [19], this process may proceed with technology as the cognitive servant (offloading the human task to the machine), as the cognitive partner (sharing the task often in complementary ways) or as the cognitive driver (drawing little on the capacity of the human analyst). In this way, the HUMMINGBIRD framework represents a truly integrative and holistic partnership between man and machine with the balance of responsibility dictated by the task and context. However, even in the latter context, it is clear that the human may offer clear benefit through inputting at various points. These are shown by dashed lines and are described below:

- (A) Reference Samples Selection – humans may usefully determine the samples that comprise the enrolment cluster for each subject. For example, a particular sample could be omitted if deemed unrepresentative or of poor quality.
- (B) Reference Sample Update – post-identification, humans could provide control over the update of a sample cluster through addition of an identified
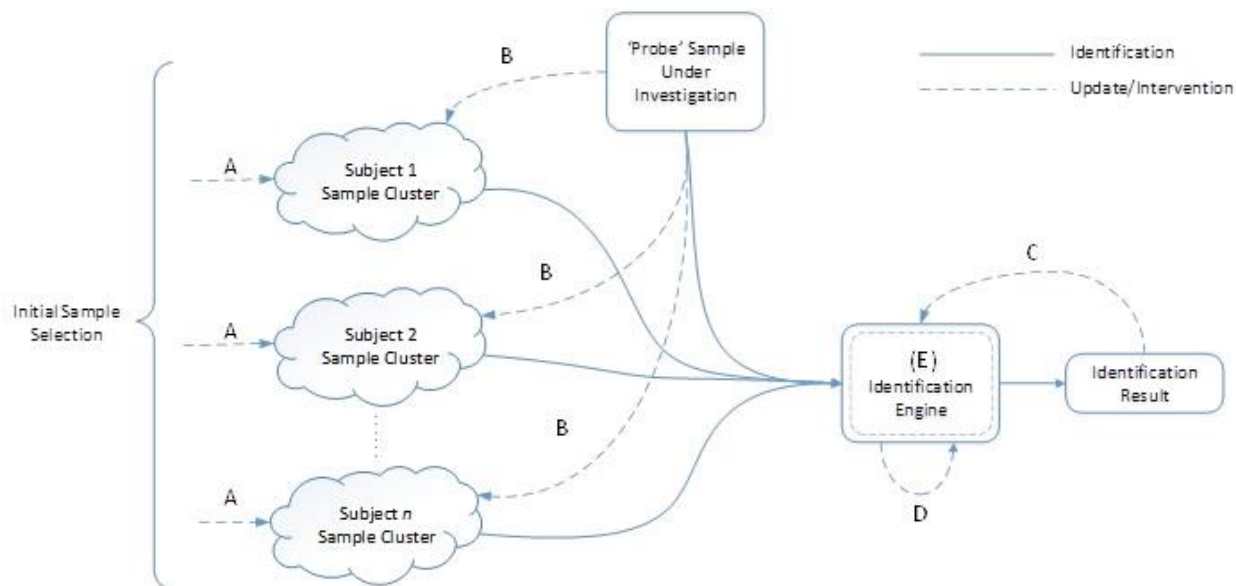
Figure 1: HUMMINGBIRD Framework

probe sample. This can be seen as a continuation of the process (A) but again provides an opportunity for the human to oversee the quality of the samples being held within the reference cluster.

- (C) Correction of Misclassified Results – the human can take the role of supervisor in a supervised training scenario, correcting the identification engine when a misclassification is deemed to have occurred. This process may involve the re-evaluation of feature or classifier topology weights to account for the error. Alternatively, it may involve re-design of the engine itself through a continuous improvement cycle.
- (D) Features under Investigation – prominent human cues used for identification can be incorporated within the identification engine, either at the point of design or through a supervised tuning process. Using the previous example, vocal distinctiveness is a cue known to affect the accuracy of voice matching when implemented by the human analysts, and this may be incorporated to enable human-weighted decisions.
- (E) The identification engine may be a machine-based solution or a human one. However, more exciting and potentially of greater probative value is a decision-making process which integrates complementary human and machine judgements, as outlined earlier in this paper.

6. Emerging Questions

In describing the HUMMINGBIRD framework, several research questions naturally emerge. Far from raising superficial or quantitative concerns, these questions draw on areas of theoretical interest in both human and machine vision. For instance, when considering the samples that should be included within a reference sample cluster, issues of canonicality and representativeness emerge [20]. Both are areas upon which the cognitive psychology literature has much to offer. Equally, there are issues of whether to include a broad variety of samples within a reference cluster, in order to support identification under challenging conditions, and these must be tensioned against issues of parsimony and practicality. These themes will be familiar to psychologists in terms of exemplar models versus prototype models of recognition (see [21]). Similarly, the consequences of these decisions will also be familiar to psychologists in terms of the potential for the sample clusters, or cognitive representations, associated with different individuals to overlap as each becomes more broadly defined. The obvious consequence is that, as the sample cluster associated with each individual is enriched by increasingly diverse samples, the capacity to differentiate between two similar individuals becomes compromised.

At an empirical level, a balance point will be required. However, the human analyst achieves this naturally through cognitive processes associated with categorical perception [22] in which similarities between two samples from the same person become minimised and differences between two samples from different people become maximised. Supervised learning is the perfect vehicle through which to simulate this human approach, and may critically enable the human to be much more than a supervisor of automated decisions.

Finally, the intriguing possibility arises as to the

desirability, or otherwise, of completely automating some biometric decisions creating what Dror and Mnookin [19] refers to as 'lights out' processes (because they can still be achieved even when the lights are turned out in the lab). Similarly, the possibility exists to adopt a completely manual approach. For the HUMMINGBIRD framework to be both dynamic and flexible, both approaches are retained as possibilities. However, clear safeguards will be required in order to protect against the weaknesses of each decision maker.

7. Conclusion

The current paper has explored the benefits of biometric fusion, within human or automated decision makers and, more excitingly, across them. A novel HUMMINGBIRD framework has been outlined which seeks to formalise this fusion approach and from this a number of research questions have emerged which demand a broad approach in their investigation. It is hoped that the HUMMINGBIRD framework will provide a catalyst for future research along these lines, propelling multidisciplinary biometrics research towards an exciting new chapter.

[1] GO Science (2015). Forensic Science and Beyond: Authenticity Provenance and Assurance. https://www.gov.uk/government/publications/forensic-science-and-beyond

[2] http://www.innocenceproject.org/causes-wrongful-conviction/eyewitness-misidentification

[3] White, D., Kemp, R.I., Jenkins, R., Matheson, M., & Burton, A.M. (2014). Passport Officers' Errors in Face Matching. PLoS ONE, August 18th, DOI: 10.1371/journal.pone.0103510

[4] Starbug (2015). I see, therefore I am… You. *Biometrics2015*, 13-15th October, London.

[5] House of Commons (2015). *Current and future uses of biometric data and technologies.* http://www.publications.parliament.uk/pa/cm201415/cmselect/cmsctech/734/73402.htm

[6] Ross, A., Nandakumar, K., & Jain, A. (2006). *Handbook of Multibiometrics,* 1st Edition, Springer Science & Business Media LLC.

[7] Kiltter, J., & Poh, N. (2008). Multibiometrics for identity authentication: Issues, benefits and challenges. *2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, 2008.*

[8] Phillips, P. J., *et al*. (2009). Overview of the multiple biometrics grand challenge. *Advances in Biometrics*. Springer Berlin Heidelberg, 705-714.

[9] Kaplan I.T., & Carvellas, T. (1965). Scanning for multiple targets. *Perceptual and Motor Skills, 21*, 239-243.

[10] Menneer, T., *et al.* (2012). Costs in searching for two targets: Dividing search across target types could improve airport security screening. *Applied Cognitive Psychology, 21(7)*, 915-932.

[11] Dror, I.E., (2015). Cognitive and Human Factors. Ch 4 of [1].

[12] Raymond, N.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175-220.

[13] Stammers, S., & Bunn, S. (2015). Unintentional Bias in Court. POSTnote 512, http://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PN-0512

[14] Stevenage, S.V., Clarke, G., & McNeill, A. (2012). The "other-accent" effect in voice recognition. *Journal of Cognitive Psychology, 24(6)*, 647-653. DOI: 10.1080/20445911.2012.675321

[15] Goggin, J.P., Thompson, C.P., Strube, G., & Simental, L.R. (1991). The Role of Language Familiarity in Voice Identification. *Memory and Cognition, 19*, 448-458.

[16] Pollack, L., Pickett, J.M., & Sumby, W.H. (1954). On the identification of speakers by voice. *Journal of the Acoustical Society of America,* 26, 403-406.

[17] Van Lanker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: Patterns and parameters. Part I: Recognition of backwards voices. *Journal of Phonetics,* 13, 19-38.

[18] Stevenage, S.V. (2015). The multiple biometrics behind a SuperIdentity. *Biometrics2015*, 13-15th October, London.

[19] Dror, I.E., & Mnookin, J.L. (2010). The use of technology in human expert domains: Challenges and risks arising from the use of automated fingerprint identification systems in forensic science. *Law, Probability and Risk, 9*, 47-67.

[20] Blanz, V., Tarr, M.J., & Bülthoff, H.H. (1999). What Object Attributes Determine Canonical Views? *Perception, 28(5)*, 575-599. DOI: 10.1068/p2897

[21] Valentine, T. (1991). A unified account of distinctiveness inversion and race in face recognition. *Quarterly Journal of Experimental Psychology, A, 43A*, 161-204. DOI: 10.1080/14640749108400966.

[22] Rosch, E., Mervis, C.B. Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 7,* 573-605.