

Document Identity, Authentication and Ownership: The Future of Biometric Verification

M.C.Fairhurst
Department of Electronics, University of Kent,
Canterbury, Kent CT2 7NT, UK
M.C.Fairhurst@ukc.ac.uk

Abstract

Document security is an increasingly important element in the multi-faceted discipline of document processing, and authentication of individual identity will play an increasingly important future role in relation to questions of document ownership, identity and confidentiality. Biometrics-based techniques are emerging as key elements in the drive to address security and confidentiality in an effective way, yet past experience suggests that there are many practical issues yet to be resolved if biometric technologies are to fulfill their potential in the document processing field. This paper addresses some aspects of biometric processing which are becoming increasing priorities, and suggests how a greater engagement of the document processing community can help to bring about refinements to existing approaches to biometric identity checking.

1. Introduction

The changing nature of both formal and informal transactions and interactions in the modern world is such that security, in its broadest sense, is an increasingly high priority in most aspects of everyday life. There are increasing demands for individuals to be directly accountable for their actions while, at the most basic level, trust and confidence in transactions requires that individual claimed identity must be authenticated by the most rigorous means possible, rather than being assumed. Only in this way, it is argued, can social cohesion be maintained, and the avoidance of crime, fraudulent use of false identity and the other undesirable aspects of increasing automation of interactions be avoided.

The question of "identity theft" is especially important, since this type of crime is among the fastest growing in the western world, and this is exacerbated by the increasing penetration of

automated interactions between clients and business and, indeed, on a direct inter-personal basis. The field of document processing has certainly not escaped the challenges of regulating and managing security, and there are many examples of situations where confidentiality, trust, and restriction of access, are essential requirements in a real-world document processing scenario.

This paper will introduce and discuss some of the important current issues which, it is argued, must be addressed if real progress is to be made in the critically important area of identity authentication. Although the relevance to the specific field of document processing of the ideas introduced will be demonstrated, such is the pervasiveness of the urgent need to check and authenticate individual identity that the thrust of the paper will be to concentrate on the generic issues which underpin the successful practical exploitation of the techniques described. Also, not surprisingly, the target for much of the work discussed is primarily in electronic document processing. However, as a starting point we can easily identify three different, though related, general areas of document processing within which a reliable identity checking system could play a vital role. These are:

- Analysing direct biometric document content

The most obvious example here is in automated bank cheque processing, where a handwritten signature is tightly bound with the document content, and where a rigorous means of authentication is essential for reliable automation of the analysis process.

- Controlling document access
Here we might envisage a situation where access to an electronic

document is to be restricted and where simple password checking is deemed insufficiently reliable. It is clear that there is great scope for introducing an access control scheme based on the use of biometric checking to regulate access to the document content.

- Forensic document inspection
The integration of document content analysis with biometric processing is just a short step away from using (at least some) biometric-related data in a forensic analysis context, where the goal is to compare document fragments or in some other way to extract information about the creator of the document. Most naturally, of course, this will involve documents which are handwritten.

The key to much of what is hinted at above lies in the developing discipline of biometrics, and it is the use of biometrics as a tool for increasing confidence in the veracity of claimed identity on which this paper will focus.

2. Biometrics and identity authentication

Biometric measurements – the measurement of attributes of an individual which help to identify that person uniquely [1] – can be drawn from a wide range of “modalities”. Some examples of common biometric modalities of current interest are:

- Facial features
- Voice characteristics
- Fingerprints
- Handwritten signature
- Iris patterns
- Hand shape
- Hand vein patterns
- Keystroke dynamics
- Odour
- Ear shape
- Gait patterns
- Retinal blood vessel patterns

Biometric measurements may be categorised as either *physiological* or *behavioural*. The first type, examples of which include iris patterns, fingerprints etc, relate to inherent physiological characteristics of an individual, while the second type, such as handwritten signatures, keystroke dynamics, gait patterns, etc, arise from activities carried out by that individual, either those which occur spontaneously or, in some cases, those which are specifically learned.

The list given above shows an extensive range of possible biometric modalities, illustrative rather than exhaustive, since others could also be envisaged. The existence of so many options can be useful, since it allows a choice to be made to support identity checking with a range of different individuals, different environments and different application domains. Although the handwritten signature has perhaps the longest history as a specifically biometric data source, many other biometric modalities, often exploiting technology originally developed primarily for other purposes (such as face or voice recognition), have gained popularity in recent years.

Although for most biometrics widespread practical exploitation has been a comparatively recent phenomenon (indeed, practical biometric implementation is an area which is still moving towards maturity), research interest in all forms of biometric measurement has a long history, and experience has clearly shown that different biometric modalities offer their own advantages and disadvantages. Similarly, an important factor in choosing an appropriate biometric for a given task is the degree of acceptability which it is accorded within the prospective user community. Although the introduction of biometric authentication can sometimes be accompanied by scepticism or even outright opposition (for example, because of poor reliability, concerns about privacy and civil liberties, and so on), there is general agreement that in many applications biometrics-based schemes offer advantages over other approaches to identity checking.

More than many areas of academic study, however, the emerging discipline of biometrics has attracted some interesting commonly encountered misconceptions.

3. Some common myths

Here are some typical biometrics myths.

1. *Biometrics is a solved problem because many commercial devices are available right now.*

Not true. You can buy many devices, but there is no device which is a universal solution for all tasks and all situations.

2. *We don't need so much choice in choosing a biometric to work with.*

OK, but this overlooks the obvious fact that biometrics measure things about people, and people are all different. Not every biometric is equally suited to every potential user.

3. *Biometrics is easy. Just pick the algorithm you want and plug it in*

Not really. Although any chosen algorithm may have a simple operation in principle, the nature of many situations where it may be used is such that a highly complex series of transactions and “negotiations” may be needed to operate even with an intrinsically simple biometric device.

4. *Don't worry about the interface. It's easy.*

Partly true, in the sense that, if you choose a well-developed device which is suitable for some simple operating environment, then you may not need to do much more to make a start. However, you'll soon find that if you really want to exploit the full potential of biometric processing, you'd really like to do more than this.

5. *Using biometrics is good because I can always trust objective biometric data*

Not recommended. Such is the nature of current technology that increasingly sophisticated techniques are available to generate data which might fool a biometric checking system. Think about synthesized speech patterns, artificially constructed face images, and so on.

The message is therefore straightforward. Biometrics may appear to be a simple concept but, for real practical viability, there are many important issues to resolve. Furthermore, there are many on-going problems on which progress still needs to be made if we are fully to realise the potential of biometrics, as will become apparent. Thus, the next section points to some current “hot topics” in biometric processing.

4. Research issues in biometrics

4.1 User characteristics in biometric interactions

If biometric identity checking is to be effectively adopted, it is essential to understand the nature of the data which is to be processed.

Trials have been conducted to collect biometric data from a cross-section of the general public, using three commercial devices which operate respectively with facial images, fingerprints, and voice samples. Details of these trials can be found in [2].

Volunteers (221 in total) were recruited, each undertaking two data collection sessions. The first session entailed enrolment on each of the systems adopted (up to three enrolment attempts per device were permitted), together with a post-enrolment verification check to ensure that the enrolment process had been successful. At a second session at least one month later a verification process was carried out using the original enrolment models. An investigation of the longer term variability of data is important but was not considered in this study.

Modality	Failure to enrol rate
Fingerprint	14.4%
Voice	1.5%
Face	27.1%

Table 1: Failure to enrol rate for each modality

Modality	Failure to verify rate
Fingerprint	2.7%
Voice	9.9%
Face	1.4%

Table 2: Failure to verify rates for each modality

Tables 1 and 2 show the “failure to enrol” rates encountered for the three modalities, and also show the verification failure rates subsequently measured. It is clear that achieving a satisfactory enrolment can be difficult and that the degree of difficulty is device dependent. It is also clear, even from such a preliminary study, that ease of enrolment does not guarantee high levels of performance accuracy subsequently.

It is interesting to explore this issue a little further, and Table 3 shows the failure to enrol rates as a function of the number of enrolment attempts allowed for the case of the fingerprint modality. This demonstrates exactly why multiple enrolment attempts are generally necessary in practice and, especially, shows how a much poorer performance would be recorded if only a single enrolment attempt were allowed. A clear message here is the illustration of the positive effect of “training during use” associated with this type of activity, and we will return to this point later.

	Failure to enrol
1 st enrolment attempt	28.5%
2 nd enrolment attempt	7.7%
3 rd enrolment attempt	2.7%

Table 3: Failure to enrol as a function of number of enrolment attempts (Fingerprint modality)

4.2 Improving performance using multi-modal biometrics

There is an increasing recognition that, in order to achieve the robustness and reliability required for many applications of biometric identity checking, no single biometric modality is likely to meet all the criteria associated with a specified task domain. Thus, increasingly, *multi-modal* biometric checking is likely to be adopted as a

practical solution which can offer two immediate benefits. First, combining evidence of identity from more than one biometric source can obviously improve the accuracy which can be achieved. Second, the availability of multiple modalities can offer a flexible solution to the problem of unreliability of any particular modality in a given situation or to the question of unacceptability of a particular modality for any particular user.

The potential to improve performance through the availability of more than one modality can be illustrated simply by returning to the data obtained from the user trials. For example, we can consider the improved performance achievable simply as a result of the availability of more than one modality, and where a weakness in one modality can be compensated by strength in another. Table 4 shows the Failure to Enrol rate and the Verification Failure rate when a combination of the fingerprint and voice biometric modalities is adopted. In the first case a Failure to Enrol is defined as a failure to complete a satisfactory enrolment process on both of the available devices, while Verification Failure refers to the situation where successful verification based on at least one of the modalities in the chosen combination has not been achieved.

The results show that a significant improvement in performance can be obtained in the multiple modality scenario. In fact in other combinations considered (for example, combining, say, fingerprint data and facial characteristics or voice and face features) it can be shown that it is possible to reduce these error rates to zero with appropriate choice of the modalities to be combined.

	Failure rate (%)
Failure to enrol	0.9
Verification failure	0.5

Table 4: Performance characteristics for dual modality scenario

In fact, there is an extensive and varied literature on multi-modal identification systems [see, for example, 3,4,5]. However, in most reported

work attention is generally focused on a multi-modal identification procedure based on a fixed set of biometrics. For instance, in [5], fingerprint and face modalities are used, while in [3] and [4] it is visual and acoustic aspects of the voice which are processed. For greater flexibility it is important that a variable set of biometrics can be accommodated, to develop a multi-modal system which can be configured according to the demands of a particular task domain.

There are, however, a large number of unresolved issues in multi-modal biometric processing, especially relating to questions about how to combine evidence from the different modalities and how to structure a multi-source classifier/verifier.

4.3 Implementation issues, data management and user-centred design

In considering the use of multi-modal biometrics in a realistic setting, such as document access and management, it is clear that there are several inter-related sources of variability which are likely to affect the required performance of the authentication system. These sources include, for example, environmental conditions, users' physiological/behavioural characteristics, users' preferences, variability of the communication channels, and so on.

Thus, there is a clear requirement for a biometric system to be able to adapt to user needs and conditions and, especially, to be able to determine and maintain an acceptable balance between confidence and convenience for its users through negotiations between information users and providers. The result of this facility is that the system becomes able subsequently to adapt quickly and efficiently to changing external conditions.

A key to a user-centred design is to be found by returning to the idea of a multi-modal processing scenario. A multi-modal system should be able to deal with situations where a user may be unwilling or, in many cases, unable to provide a certain biometric, or where a preferred biometric is insufficiently accurate for adoption. In this sense, according to the characteristics of the current situation, a multi-modal system has to adapt to user and environment needs and use the

most appropriate set of biometrics available.

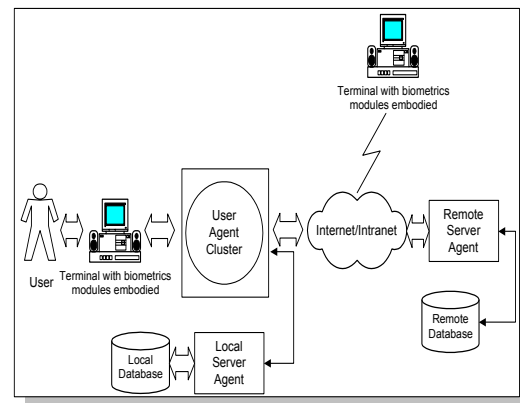


Figure 1: The structure of the IAMBIC system

An interesting approach to the management of such a system is to implement an architecture which uses agent technology to realise the biometric-based processing. Intelligent autonomous agents [6,7] and multi-agent systems represent a research field of much current interest with a variety of applications. Agents are (sub-)systems which interact intelligently with their environment, and are capable of autonomous action. In addition they are flexible in their responses, pro-active in exploiting opportunities and seeking goals, and "social" in their interactions with other agents. They may also exhibit other valuable properties such as adaptability or mobility.

Multi-agent systems are implemented as a group of several interacting agents and are well suited to situations where multiple perspectives of a problem-solving situation exist. Types of interaction that may best be suited to biometric security involve co-operation, co-ordination and negotiation between agents. The needs of the information provider for establishing sufficient trust in the user may have to be balanced with the confidentiality of the user's biometric information and his ease of use of the system. A balance may need to be struck for each service, transaction or session and may even be dynamically modified during use.

Such appropriate and highly flexible approaches to biometric system implementation have often been overlooked, but consideration of this issue is essential if biometric systems are to be effective and reliable in practical applications.

human voice, fingerprint, signatures and facial images offer opportunities for compromise and state-of-the art techniques will need to be studied in objective and subjective tests of biometric recognition vulnerabilities for both humans and machines.

Context Recognition: Similarly, recognition of *inter alia* environmental, biological, semantic and emotional context can provide clues to better biometric protection. These can be investigated as a means to establish the context for biometric transactions in order to increase robustness in decision making. Liveness detection, authorship analysis of text, and related techniques may be identified as suitable tools with which to address this issue.

Context Integration: Strategies can be developed for biometric algorithms which can integrate context sensitivity to provide counter-countermeasures to purposeful attack. These will include multi-modal strategies (see above) and challenge-response methods. Interestingly, against this background it may also be valuable to report some attempts to understand human behaviour in biometric processing.

Signatures and signing styles can differ significantly, both within samples from the same signers, but self-evidently to a very large degree across a population of signers, and the susceptibility of a signature to false imitation is intuitively a function of the nature of the signature itself.

In order to investigate signature perception, in a further experiment subjects were asked to view a range of signatures of varying perceived “complexity”, some of which were genuine samples and some of which were forgeries (generated in a separate experiment with a disjoint set of subjects who produced the imitations from a visual inspection of a genuine sample). In total each subject viewed 10 genuine and 10 forged samples from each of the five target groups. Each subject was asked simply to classify each sample as “genuine” or “forgery”, in comparison with a genuine sample which was in view simultaneously, as would be the case, for example, in checking a signature against a

Signature Group	Avg. Perceived Complexity	Total error (%)	FAR (%)	FRR (%)
A	1.8	33.1	2.1	30.9
B	4.8	27.5	2.1	25.4
C	8.2	21.9	2.9	19.0

“model” written on some reference document.

Table 5 shows the results of this experiment. In order to give clearer emphasis to the degree of distinguishability in the judgments made by participants, the signature samples which were rated the least complex and the most complex were selected, together with the signature ranked in the middle of the overall ordering. These are designated sample A (least complex), sample B (middle) and sample C (most complex) in the results reported. The Table shows the average perceived complexity value, the total number of authentication errors made, and the split between false acceptance rate of forged signatures (FAR) and false rejection rate (FRR) of genuine samples.

An intuitive assessment of the relation between perceived complexity and the likelihood of errors in judging sample authenticity can lead to two (essentially opposing) hypotheses. On the one hand, it could be predicted that low complexity leads to ease of imitation and therefore potentially higher FAR. On the other hand, it might be supposed that higher complexity makes imitation more difficult, and therefore more flexibility in assessment, leading to more errors in perceived authenticity, might be expected. It is the latter hypothesis which appears to be supported by the results of our study, and this information can be most useful in developing strategies to maximise the effectiveness of biometric processing.

4.6 Testing and evaluation

Table 5: Error rate performance and complexity assessments

The importance of appropriate testing and evaluation of biometric systems is often overlooked. However, inappropriate testing strategies have often led to misleading conclusions about the likely impact of biometric processing, and this is an area to which increasing significance is being accorded.

It is essential that the conditions under which testing takes place are strictly and uniformly controlled, that the test protocol adopted is carefully defined and consistently applied, and that the test participants are selected and briefed appropriately. Only then can the results of an evaluation trial be interpreted in a meaningful way. Best practice guidelines are beginning to emerge and, though testing scenarios are often not considered, this is an area where increased rigour in the future will certainly pay great dividends.

5. Document processing and biometrics

The review and discussion outlined above shows that the field of biometrics is a vibrant and still-maturing discipline, but also that a degree of optimism is appropriate in considering the way in which the viability and effectiveness of biometric data processing might be increased in the future.

In parallel, however, the discussion also suggests some further challenges which remain for both the biometrics and the document processing communities if the two areas of interest are to be effectively integrated to their mutual benefit. In conclusion, therefore, it is appropriate to define some challenges for biometrics in document processing which still need to be confronted. Here are some suggestions:

- **Offline biometric processing:** This is particularly an issue in applications such as automatic signature checking on documents. In many applications (for example, automated bank cheque processing) the biometric (signature) data must be extracted offline, since the document is processed remotely from its point of creation. This limits the nature of the features which can be extracted from the biometric sample. It is generally recognised that most rapid progress has been made in on-line processing, where the more

information-rich dynamic as well as static features are available, and reliable off-line signature verification remains a significant challenge.

- **Document structure:** Document creation and the specification and implementation of useful document models is a fertile and diverse research area. However, if biometric-based security is to be incorporated into a document processing scenario, then the choice of an appropriate document model can be critical in supporting this. Balancing biometric security requirements against competing issues concerning implementation flexibility and generality raises important questions for the future.
- **Improved document access control:** Of course, at the most basic level, biometric-related document management can simply relate document access to user identification based on an interface which is linked through a conventional biometric checking device. Although simple in principle, there remain very many questions about how to optimise such systems and, indeed, how to achieve levels of performance which make the process viable in practical situations. This, of course, is the main focus of much of the preceding discussion in this paper, and it is clear that there are many challenges ahead. Notable among these, however, are issues relating to the design of intelligent adaptive interfaces which can offer the opportunity for smoother integration of the biometric checking procedures, providing an optimal match between specific user characteristics and the generic capabilities of a particular biometric modality.
- **Biometrics-linked encryption:** Encryption is the accepted means of providing security to electronic document contents but, of course, encryption protects users only after the point at which the user has launched an application. Biometric identity checking can enhance the overall level of document security offered by

ensuring that users are appropriately authenticated both pre- and post- data transfer. An ultimate objective might therefore be to develop a means of linking the encryption/decryption of document content directly to user identity as encapsulated in biometric measurements. The benefits and potential pitfalls of such a process are readily evident, but this is an area of great current interest.

6. Conclusion

Questions about document ownership, about regulation of access to documents, about authorising document-based transactions and about tracking document history will increasingly raise issues about how to establish the identity of individuals in a reliable and robust way. The increasing penetration of biometrics as a means of providing techniques to achieve these objectives both for physical documents (e.g. in bank cheque processing) or, especially, in respect of electronic documents, suggests that the document processing community should engage with current trends in biometrics in an increasingly widespread way.

This paper has introduced aspects of biometric processing which are likely to lead to significant developments, and has demonstrated some approaches to the exploitation of biometric techniques in some important potential document processing applications.

As with all emerging techniques, biometrics offer both potential solutions to some difficult problems, but also real challenges. The fundamental aim of this paper is both to stimulate greater awareness of some current important issues, but also to suggest some promising directions which might constructively address these challenges.

7. References

1. A. Jain, R. Bolle, S. Pankanti (Eds.), Biometrics: personal identification in networked society, Kluwer Academic Publishers, 1999
2. M.C. Fairhurst, J. George, F. Deravi, Scenario-based data collection trials for the evaluation of multi-modal biometric processing: a preliminary report, Proc. KES2002,
3. Chibelushi CC, Deravi F, Mason JSD (1999) Adaptive Classifier Integration for Robust Pattern Recognition. IEEE Transactions of Systems, Man and Cybernetics - Part B: Cybernetics. Volume 29, No. 6, pp 902-907
4. Su Q, Silsbee PL (1996) Robust Audiovisual Integration Using Semicontinuous Hidden Markov Models. Proc. of the Fourth International Conference on Spoken Language Processing, Vol. 1, pp. 42-45
5. Hong L, Jain A (1997) Integrating Faces and Fingerprints for Personal Identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(12): 1295-1307
6. Wooldridge M, Jennings NR (1995) Intelligent Agents: Theory and Practice. The Knowledge Engineering Review, 10(2), pp 115-152
7. Jennings NR, Sycara K, Wooldridge M (1998) A Roadmap of Agent Research and Development. Autonomous Agents and Multi-Agent Systems, Kluwer Academic Publishers, Vol 1, pp 275-306
8. S. M. Brown, E. Santos Jr., and S. B. Banks. (1998) Utility theory-based user models for intelligent interface agents. In Lecture Notes in Computer Science 1418: Advances in Artificial Intelligence -- AI '98, pp. 378-392, Springer-Verlag

Acknowledgement:

Much of the work reported in this paper arises from collaboration with other colleagues in the Department of Electronics at the University of Kent. The author gratefully acknowledges their contribution and, in particular, would like to thank Dr.F.Deravi, Dr.K.Sirlantzis, Dr. S.Hoque, Mr. N.Mavity, Ms J.George and Ms E.Kaplani.