Security Analysis for Biometric Data in ID Documents

S. Schimke[a], S. Kiltz[a], C. Vielhauer[a], T. Kalker[b]
[a]Otto-von-Guericke University Magdeburg, Universitätsplatz 2, D-39106 Magdeburg
[b]Hewlett-Packard Labs, Palo Alto, 1501 Page Mill Road, CA 94304, U.S.A.

ABSTRACT

In this paper we analyze chances and challenges with respect to the security of using biometrics in ID documents. We identify goals for ID documents, set by national and international authorities, and discuss the degree of security, which is obtainable with the inclusion of biometric into documents like passports. Starting from classical techniques for manual authentication of ID card holders, we expand our view towards automatic methods based on biometrics. We do so by reviewing different human biometric attributes by modality, as well as by discussing possible techniques for storing and handling the particular biometric data on the document. Further, we explore possible vulnerabilities of potential biometric passport systems. Based on the findings of that discussion we will expand upon two exemplary approaches for including digital biometric data in the context of ID documents and present potential risks attack scenarios along with technical aspects such as capacity and robustness.

Keywords: Passport, ID document, biometrics, security, barcode, RFID, digital watermarks, hologram

1. INTRODUCTION

Biometric user authentication techniques have evoked an enormous interest by science, industry and society in the recent past and significant improvements of biometric recognition techniques can be found for methods based on different physiological and behavioral modalities. Driven by the possibility to provide methods for automated determination or confirmation of the identity of subjects based on their physiological and behavioral traits, these technologies are highly interesting for identity verification at international borders. Consequently, it has been decided by governmental institutions in Europe and the U.S. to include digital biometric data in future ID documents. However, this intention brings some new security risks, which are discussed in this article.

In order to identify these problems, it is necessary start with an overview of the most adequate biometric features, which appear interesting in this context. Further, the technical aspects of data storage techniques have to be reviewed, and security goals for this particular application have to be defined. The two main security aspects for classical ID documents are the ability to detect tampering and to prove authenticity. It will be shown that in the context of the extension by biometric features, security requirements also need to be extended to confidentiality, with the inclusion of non-overt biometrics. The goal of this paper is to elaborate on the problem description, expose vulnerabilities of various approaches and to identify future research directions.

2. SECURITY GOALS OF ID DOCUMENTS

For ID documents, e.g. passports, there are several requirements regarding security aspects. First, it must be guaranteed that only legitimate instances (licensed authorities) are able to produce the document (*document authenticity* or *unforgeability*). This means that no illegitimate subject should be able to produce an ID document in a way that the forgery can't be identified as forgery. Second, there has to be a verifiable relation between the ID document and its legitimate holder (*non-transferability*), i.e. no subject **A** is able to present subject **B**'s ID document as its own. Third, it must be possible to recognize manipulations on ID documents (*integrity*). This means subject **A** has to be prevented from manipulating **B**'s ID document in such a manner that **A** creates an incorrect relation between itself and **B**'s ID document. Note, that there is a difference between document authenticity and integrity in ID document scenarios.

3. CLASSICAL TECHNIQUES

In the past, security requirements as described in section 2 were addressed for example by using special materials and applying proprietary production engineering. To achieve the security criterion of *non-transferability*, most ID documents contain a photograph and an image of the handwritten signature of the legitimate holder and in some cases some information on the holder's appearance, like body height or eye color. The security criterion of *authenticity* is often obtained through the appliance of hard-to-forge production technologies, e.g. embedding watermarks into the document paper, holograms, security foils, micro script and special inks. Hitherto the *integrity* of an ID documents has been

Security, Steganography, and Watermarking of Multimedia Contents VII, edited by Edward J. Delp III, Ping Wah Wong, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 5681 © 2005 SPIE and IS&T · 0277-786X/05/$15

gained by stamping and clamping the photograph and more recently by laminating the document. Stamps and clamps are intended to prohibit replacement of the photograph with one of another person. By laminating the documents, subsequent modifications are made impossible. Using special embossing techniques in the laminating process ensures authenticity as well.[6, 32]

Figure 1: Example of a German passport.

4. BIOMETRICS

The term *biometrics* in this paper refers to the science of measuring identifying features or attributes of human beings. We distinguish two approaches; passive and active schemes. In passive biometrics, physiological characteristics of human subjects are examined. Examples of passive biometrics are fingerprint, iris or face recognition. In contrast to the passive biometrics, for active biometrics behavioral characteristics of persons are considered. Speech, gait, handwriting, keyboard typing or lip movements are examples of these behavioral biometrics.

A human attribute is feasible for biometric authentication if it is stable for a subject and if the subject can be discriminated by that attribute. *Stability* in this case means that the attribute varies little in between two measurements, especially over longer periods of time. A biometric attribute has *discriminatory* power, if most persons have a different characteristic for that attribute. An example for a non-stable and non-discriminatory attribute is the body weight – it tends to change for a person and in a large population many subjects have the same weight. In contrast, a feasible attribute is the ridge pattern on the skin of a fingertip, also known as a fingerprint. The fingerprint is known to stay the same across the lifetime of a person and it is hard to find different subjects with the same fingerprint pattern.[*] Finally, for convenient and automatic application the biometric attribute has to be quickly acquirable; to be appropriate for security relevant scenarios, it has to be as demanding as possible to fake the attribute.

A biometric system generally works in following five main steps. First, the biometric attribute of a person has to be measured. In case of fingerprint biometrics, this is performed by using a sensor to scan the fingertip surface. From this raw sensor data in a second step, special characteristic feature data are extracted and stored as the reference of the respective subject. This could be for example a special minutiae point in the fingerprint structure. These two steps are referred to as *enrollment* (see left part of Figure 2). In the case of authentication, the stored reference data is compared with actually acquired data. Here, as steps three and four, the person's attribute is scanned again and its features are extracted. The actual main part of authentication is the fifth step: *comparison*. Here, using a similarity or distance function, a measure of similarity between the reference and the acquired data is computed and a decision about their equality is the derived. Two approaches for authentication can be distinguished. By comparing the acquired data with reference data of one subject a *verification* is performed (1:1 comparison, see right part of Figure 2), whereas by comparing the acquired data with reference data collected from a number of people an *identification* is done (1:n comparison). Instead of only storing feature template data during the enrollment, in some systems the entire raw data is saved. The advantage of this method is to obtain interoperability between different biometric systems, which may use non-standardized kinds of templates. The technical disadvantage is of course the higher demand on storage space and the possible abuse of the stored data.

To rate the quality of biometric systems often the probability of errors is considered. Four of the most important kinds of error rates are the *false acceptance rate* (FAR), the *false rejection rate* (FRR), the *failure to enroll* rate (FTE) and the *failure to acquire* rate (FTA). The FAR is the probability for a non-legitimate subject to be wrongly accepted as a legitimate user by the system. Analogous the FRR is the probability for legitimate subject to be rejected unjustified. The FTE statistically represents the proportion of the population with a non-adequate forming of the biometric attribute.

[*] New criminological investigations raise suspicion regarding the uniqueness of fingerprint patterns, but nevertheless until now the fingerprint is adequate juristic evidence at the court. [21, 22, 47]

Complementary to the FTE is FTA, which covers those cases where for an already enrolled subject the acquisition of the biometric attribute is (temporarily) not possible as result of e.g. lesions or defilements.

In the following we will give a short survey about several selected biometric attributes: fingerprint, face, iris and handwriting, which appear relevant for security scenarios in context of ID documents. We choose fingerprint and handwritten signature biometrics for closer analysis, since both these attributes are already used in most passport documents. Fingerprints and iris biometrics are also included since these are well researched.

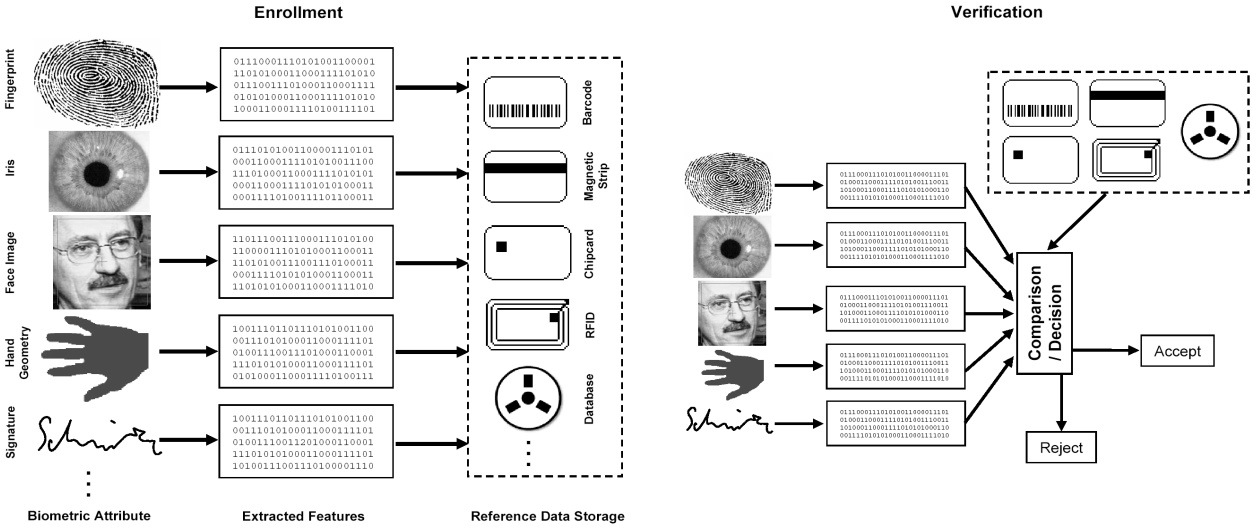

Figure 2: General biometric processing steps: enrollment and verification.

4.1 Fingerprint

One of the biometric attributes that have been studied intensively is the human fingerprint. The fingerprint is the unique skin structure of fingertips. As a phenotypic biological feature it is unique, even for identical twins. The characteristic formation of the fingerprint normally doesn't change over a person's life span.

For automatic processing, the fingerprint is taken using special sensor devices. In most cases the raw data taken is a gray scale image. To compare fingerprint images, a set of distinctive pieces of information is extracted from the gray scale images. These extracted features can be positions or configurations of ridge-lines, crossing, bifurcation and ending points, referred to as minutiae.

For storage of complete fingerprint image data, up to 250 Kbytes is needed. Using compression like WSQ can reduce this amount of space[*]. By saving only extracted minutiae information, a reduction to a magnitude of one Kbyte is possible.[33] A disadvantage of storing only extracted feature data can be the lack of interoperability between biometric systems from different vendors, if these systems use different types of feature data. This problem was addressed by defining standardized formats, i.e. *ISO/IEC JTC1 SC37 19794-2*[†] or *DIN V 6400*[‡].

The main technical problems with using fingerprint biometrics for automatic authentication purposes are the error rates. About 2% of the population has a fingertip surface with insufficient ridge/valley forming for extracting an adequate amount of feature information (2% FTE).[34] Beyond these 2% FTE the failure to acquire rate is higher in reality due to temporary lesions of defilements of fingertip skin. Recent evaluation activities such as the third *Fingerprint Verification Contest* (FVC 2004) appear to confirm this dimension of error characteristics with reported *equal error rate* (EER) for FAR and FRR of 2.07% for the best algorithm, amongst other performance indicators.[14]

Of extreme importance for the reliability of fingerprint based authentication systems in security environments is the fact that some common scanner types can be circumvented by finger surface dummies.[28, 40] This attack works by capturing a fingerprint image and producing a rubber mask for the fingertip from digital images. To acquire the finger-

[*] Wavelet Scalar Quantization (WSQ) is the standard compression algorithm for greyscale fingerprint images used by FBI. The default compression ratio for 8-bit images is 0.75 bit/pixel.[27]

[†] ISO/IEC 19792-2 – Biometric Data Interchange Formats-Part 2: Finger Minutiae.

[‡] DIN V 66400 – Finger Minutiae Encoding Format and Parameters for On-Card Matching. (DIN is the German standardization authority.)

print image, it can be taken from polished surfaces such as glass or from a biometric storage that also contains raw data. But even from minutiae data, specialized algorithms can generate a complete fingerprint image.[7, 31] To defeat these kinds of attacks, the discipline of liveness detection has emerged in the past years and a number of approaches have been presented for this purpose.[38]

4.2 Face

Face recognition has the great advantage of not requiring any sort of contact, so there are no hygienic concerns. The biometric samples here are typically taken as 2D images of the frontal section of the face, using one or more digital cameras. Typical technical approaches to achieve this recognition include geometrical, eigenfaces, template and graph matching, neural networks and Hidden Markov Models, or a combination of those.[42] Although there are significant advantages to this technology, mainly due to the overt characteristics of face images and the user acceptability, the recognition accuracy has shown to be rather inaccurate in practice. The inaccuracy can be explained by a high sensitivity to environmental conditions such as lighting and image background, but also due to changes in the appearance of a face with regards to hairstyle, beard and glasses for example. Error rates determined during evaluation of commercially available face recognition rates have shown FRR in the range of 5% at a FAR level of 10%[36] in person verification scenarios, similar to those imaginable for automated ID document verification. In identification scenarios, for example useful for video surveillance applications, face recognition systems have been reported to have shown very low correct identification rates between 2% and 36%.[4]

Despite the rather poor recognition accuracy, face recognition appears to be among the most interesting biometrics for user authentication in ID document scenarios, simply because of the simple image acquisition and the intuitive concept of comparing face images to images included in the document. Consequently, standards for the layout of facial images, as well as digital storage formats are currently developed, for example by the ICAO (*International Civil Aviation Organization*), see *ISO/IEC JTC1 SC37 19794-5*[*]. This standard recommends a storage capacity of at least 11 KB for a full frontal face image.

4.3 Iris

Iris based biometric recognition is based on the concept of comparing iris images by decomposition into Iris patterns. In practice, one predominant technique is based on the quadrant mapping of the complex phasors of 2D wavelets.[9] Using iris characteristics is said to be one of the most accurate biometric identification method, i.e. the error rates FAR is reported to be zero and FRR is very low, which has been elaborated in experiments with more than 2 million iris comparisons.[10, 33]

To this date it is very complex to forge the iris, but there are some recent research approaches in this direction.[29] Similar to the problem of forgeries of fingerprints discussed earlier, such kind of attacks are expected to be defeated by methods like infrared sensors and/or liveness detection. Although Iris recognition can be classified as highly accurate and has already been successfully applied in some border control applications, problems have been reported in the handling of image cameras, which may lead to an increase in FTA rates.[40] However reliable quantitative figures for these aspects could not be determined by the authors. Further, privacy concerns have to be considered, as individuals may fear detection of states of illness or narcotization from Iris images.

With respect to the space needed for iris templates, the IrisCode representation allow for very compact storage by a 2048 bit representation of the coded phasor quadrants.[9]

4.4 Handwriting / Signature

The handwritten signature is often used as a kind of authentication in the real life. For example, many people in daily life sign contracts, credit card receipts and similar documents. This inclusion of an explicit expression of intention is a great advantage of handwritten signatures over other biometrics. Furthermore, many existing ID documents today (such as passports issued by many countries) already contain an image of the handwritten signature as one authentication information. In the past years a wide number of approaches to signature verification have been published, some of which require offline input data (i.e. scanned images of signatures), others utilize online data, acquired from digitizer tablets. Although in general signature verification can be considered as one of the less accurate techniques, recently achievements towards higher recognition accuracy have been reported particularly in the field of online recognition. In the first International Signature Verification contest 2004 for example, EER of slightly above 2% have been reported for the most accurate algorithms.[46] However, it has been shown, that error significantly increase (by the order of one magnitude), if signature verification systems are exposed to skilled forgeries.[45]

[*] ISO/IEC 19792-5 – Biometric Data Interchange Formats-Part 5: Face Image Data.

The technical approaches for online signature recognition can be classified into function based approaches, where the similarity between multiple signals is measured based on various features and statistical techniques. The later category extracts statistical properties from the handwriting signals and compares them during the verification process. While for the first category, reference signals need to be stored as reference templates, storage requirements are in the range of several Kbytes, whereas statistical models require only a few hundred Bytes. Biometric hashes for example may achieve template sizes of 200-300 Bytes and have shown EER of down to 2%, depending on the digitizer tablet characteristics.

4.5 Summary

For security application the most viable solutions appear to be iris and fingerprint recognition with respect to recognition accuracy, particularly for automated user authentication in scenarios with little or no manual observation. Both methods have a level of inconvenience, but this may be manageable. Forgeries can be prevented by future technology by supplying sufficient liveness detection tools.

On the other hand, face recognition techniques can be used as decision support tools for manual authentication scenarios, like observed border control. While the error rates that are achieved today are too high for automated user verification, the determination of some similarity score may proof as being helpful to optimize authentication processes in some applications.

Signature Verification may be useful in applications where users can be considered as cooperative in the authentication process and user activity is required as a declaration of intention. In the application scenarios of ID cards, this does not seem to be necessary for the primary goal of verifying the identity of subjects, with or without their explicit consent. However, it seems plausible that ID documents serve a secondary purpose, e.g. authorization of transactions like check cashing. Here, the explicit consent may be a desirable feature and this could be achieved by inclusion of signature biometrics in ID documents

5. BIOMETRIC REFERENCE DATA STORAGE

As discussed in Section 4, in a biometric scenario the reference data of a biometric attribute (e.g. iris or fingerprint) needs to be saved to be available for later comparison processes. For storage of this information, database systems or mobile memory units can be used. In this section we describe different storage approaches and discuss their advantages and disadvantages regarding capacity and security aspects.

5.1 Photograph

The most common way to store biometric information in a passport is to print an image of the biometric attribute. Most ID documents contain face photographs of the holder and in some countries additionally images of one or more fingerprints are included. The advantage is that a human controller is able to check manually the relation between the ID document and the person presenting it. For the automation of passport inspection, the image (face or fingerprint) in the document and the life image both have to be digitized. The scanned image and an image actually taken image are compared as described in section 4.

For face biometrics an evaluation study of automatic recognition using different kinds of photographs has been performed in *BioP I* (biometrics in passports).[2, 5] Face recognition systems from different vendors were tested with photographs of different types and qualities in ID documents and with different poses (frontal and half side-face). The error rates in the *BioP I* study were too high for real world usage. One conclusion of the study is the recommendation to use digitally stored templates instead of photographs for automatic face recognition. The stated reason was that photographs change over the time through scratches or kinks, while digital stored data is resistant against aging.

An advantage of image-based storage of biometric data over digital storage of template data is the higher degree of interoperability between systems of different vendors, at least for as long as consistent standards for biometric data do not exist.

5.2 OCR Font

Another already widely used method for storing data is OCR (*optical character recognition*) of text. This is textual representation of information, using a special font type that is optimized for scanning and text recognition. The ICAO standardized a MRZ (*machine readable zone*) for travel documents and a number of issuers of these documents have already implemented this zone. Following *ICAO standard 9303*, the MRZ of most actual travel documents consists of two lines with each 44 OCR readable characters and can encode information, such as name, nationality or sex of the passport holder.

Due to the limited information density of text* and the limitations of accuracy in OCR recognition, this kind of data storage is only suitable for saving short feature extracts from biometric attributes, hash values or similar brief data. The great advantage of OCR text over other digital storage methods is the human readability of the data content, because any modification in the digital content will consequently result in a change of the visible text on the document.

5.3 Barcode

A different method for data storage using printing techniques is the barcode as known from product labels. Barcodes are automated recognizable graphical structures consisting of black and white areas. Mainly two different kinds of barcodes are used, one-dimensional and two-dimensional ones.

Because of the small obtainable data density of one-dimensional barcodes, we focus our discussion on two-dimensional types. The most important type of 2D-barcodes is the PDF417 (*portable data file*) as defined in *ISO/IEC 15438*. It has a maximal data density of 300 bytes per square inch and can store up to 2000 bytes†. Other types of 2D-barcodes are DataMatrix (*ISO/IEC 16022*), MaxiCode (*ANSI/AIM BC10-ISS*) or Aztec Code (*ANSI/AIM BC13-ISS*).

Different approaches using 2D-barcodes for storing biometric information have been presented[1, 37] and some countries have introduced or are currently introducing biometrics for passports or ID documents using 2D-barcodes. Examples of these countries are *Bosnia-Herzegovina*, *Nigeria* or *Guatemala*.[34]

5.4 Magnetic Stripes

A magnetic strip is a band of magnetic metal-oxide, as used to store data for example on credit cards. A magnetic strip is read out by physical contact and swiping past a reading head. The storage capacity of magnetic strip cards following *ISO 7811* standard is 1288 bits, distributed over three data tracks.

The data on a magnetic strip can be destroyed by magnetic fields, so magnetic strip cards have to be handled carefully. Furthermore, manipulations by overwriting of saved data on a magnetic strip can't be prevented, although some kinds of modifications may be detected by cryptographic binding of data on the magnet strip to other information on the document. However even in this case, this may impose a limiting factor in case of ID documents, as denial-of-service attacks may not be prevented. The advantage of magnetic strips over data storage using integrated circuits is the low unit price. Because of its limited life cycle and its vulnerability to failures and modifications, magnetic strips should not be used for long-term ID documents.

5.5 Integrated Circuits

Another approach for data storage is the usage of ICs (*integrated circuits*), e.g. ROM, EEPROM or non-volatile RAM. In the context of ID documents, two different relevant types of ICs can be distinguished – contact based and contactless ICs. The contact based ICs can appear in the form of smartcards. The contactless ICs are often referred to as RFID (*radio frequency identification*) transponder. Both types of these ICs can have memory-only functionality or advanced processing capabilities. All these types have their advantages and disadvantages, which will be discussed in the following.

5.5.1 Contact based ICs

Contact based ICs in the shape of a smartcard are standardized in *ISO/IEC 7816* and nowadays are in wide use, e.g. as card for payphones, security token for authentication purposes and key storage for digital signatures or issued by healts insurance funds to their customers. The ICs on these cards differ strongly in to their memory capacity, processing power and their security features against physical tampering. Modern contact based smart cards can have several hundred kilobytes of EEPROM. Equipped with adequate security logic, they are able to prevent unauthorized subjects from reading or changing memory contents. Furthermore actual ICs have the processing power to perform strong cryptography.

Since most contact based ICs need a kind of plastic card, where they are mounted, the use of this technology in classical travel documents is limited. As an advantage of contact based ICs over RFID transponders, the former cannot be read unnoticeably and the communication between the IC and the reader is magnitudes harder to eavesdrop than the radio frequency communication of RFID transponders.[25] Furthermore, the contact based data communication is more robust against jamming.

5.5.2 RFID Transponder

Most of the RFID transponders used nowadays are passive ones.[13] This means that they don't have a power supply of their own and obtain the necessary electrical energy via induction from the reader device. At this moment, mainly three

---

* Using an alphabet $A = \{\texttt{A-Z}, \texttt{0-9}\}$ to save digital data in a textual way, the information content of one character is ~5.17 bit.

† **http://de.wikipedia.org/wiki/PDF417**

different types of transponders are distinguished regarding their reading distance – close, remote and vicinity coupling[*]. The variety of contactless ICs is fairly great. It spans from low-end transponders with a size of 0.4mm×0.4mm (incl. an antenna) and only being able to transmit a stored 128-bit ID[15] to smart transponder chips one the other end of spectrum. Chips of the later category have the computational power to perform for example advanced cryptographic algorithms and possess a memory space of several Kbytes. For example, some modern contactless ICs are equipped with more than 100 kilobytes ROM, five kilobytes RAM and up to 72 kilobytes EEPROM, and have the computational capability to perform strong cryptography like RSA, 3DES and ECC.[20, 35] For these types of contactless ICs high secure access data access control mechanisms possible, to protect the data from unauthorized.

Because of the small and flat size, contactless ICs along with their necessary antenna are able to be embedded in actual passports. RFID transponders have the general problem of contactless communicating components – eavesdropping of the data exchange is relatively easy.[12] Furthermore, if no access control is implemented, within the standardized distance, the transponders can be read unnoticed. The first problem can be solved by transmitting only encrypted data. The latter problem is mentioned by the ICAO, who advises to embed electronically shielding material (e.g. aluminum foil) in the passport cover, so that the IC will not be readable while the passport is closed.[18] Even the selection of RFID standards with short reading distance can limit the risk of unnoticed data access.

5.6 Laser Hologram

Holograms[†] are often used to guarantee the authenticity of products like medicament or software bundles as well as banknotes or Id documents, because of the complexity of unauthorized reproduction. Holograms can even be used for data storage, as defined in *ISO 11694*. There are different smartcard-like products on the market with storage capacity of more than 1 MB.[26]

5.7 Steganography

Although steganographic techniques have been discussed in context of biometrics in ID documents[16], techniques of this category do not seem to be appropriate, as the primary goal here is information hiding in digital data with no demand for robustness. Traditionally in steganography, the goal is to hide a secret message, which is transmitted between two or more communication partners in a way to ensure confidentiality and unobservability. The more significant goals in context of id documents are integrity and authenticity and here techniques derived from steganography are digital watermarks, as will be outlined in the following subsection.

5.8 Digital Watermark

Techniques for embedding information as digital watermarks have been widely explored and a great variety of embedding and retrieval techniques have been presented for image as well as other digital media.[8] However in context of ID documents, there are problems with watermarking techniques. The main problems to be mentioned are firstly that a watermark per se does not provide any real security, as the embedding retrieval processes are solely depending on the key and algorithm chosen. Secondly, the data capacity is quite limited due to poor SNR ratios and, particularly in the case of ID documents, due to the rather small area available on the document.

Nevertheless, first approaches to include biometric data in ID cards using watermarking have been published. For example Hologram Watermarks have been suggested to store references of handwritten signatures in the facial image.[11] Although a media specific compression technique has been applied to the data, this approach clearly unveils the capacity limitations of reasonably robust watermarks and consequently, severe difficulties can be expected in practical applications for most biometrics, as most reference data sizes exceed several hundreds of bytes.

However digital watermarks appear to be qualified to cryptographically bind visible representations to of the ID document to biometric data stored by other techniques (e.g. MRZ or bar codes). This can be achieved for example by cryptographic hashes or message authentication codes (MAC), calculated for the biometric data and embedded in the document using watermarking techniques. The conceptual feasibility has been shown[37], where 2D barcodes have been suggested for storage of the biometrics in combination with integrity (robust) and copy protection (fragile) watermarks.

5.9 Central Database

To overcome storage space limitation and some of the security issues inherent to the RFID-technology, the movement of critical data from the transponder into the backend of a host system needs to be considered. This of course has the advantage that the data stored within the transponder can be reduced to a unique device ID that can be looked up in a

[*] Close coupling transponders are defined in *ISO 10536* and should work within a range of 1cm. *ISO 14443* defines the communication range at 15cm. The vicinity coupling transponders following *ISO 15693* work within a range of round about 1.5m.

[†] Sometimes, in this context the terms “identigram” or “kinegram” are used.

central database that forms the backend of such an RFID-system. While this significantly reduces the risk of critical biometric data being compromised or accessed due to being sent over the air or the electrical contacts, it raises questions with regards to that central database.

For our context of ID cards and passports this would mean that every party taking part in the system has to be granted access to such a database. That includes at least law enforcement officials and border control authorities of every single country taking part in such a system. This scenario would bring along a multitude of severe problems, only some of them can be mentioned here. Besides several technical issues, which would have to be solved in such a large-scale scenario and the requirement for standards for the access protocols and data formats, there are legal implications. This approach would be a serious legal challenge for those countries that by law do not allow for such a central database to be put up and maintained, due to privacy regulations. Furthermore, any centralized data collection implies the risk of potential misuse by unlawful personal having access to it. As in the case of biometric-enabled passports, such a centralized database would be required to be accessible by government offices in many different countries with different levels if bilateral trust, such a centralized approach does not appear realistic at the current point in time.

5.10 Combinations

To increase the level of security, it is possible to combine some of the technical methods for data storage in ID documents mentioned earlier in this section. For example, a passport containing a RFID-chip is more secure against attacks, if the chip is hard to be separated from the paper holding it. Here, for instance a cryptographic checksum or MAC of the data included digitally can be included in both the RFID-chip and additionally encoded in the passport document using an additional technique such as barcode or digital watermark. This would undeniably link the RFID-chip to the document containing it, i.e. ensure integrity between the digital and the visible content.

5.11 Summary

In this section, we have briefly outlined the most significant storage techniques, which may be used in context of biometric ID documents. While each of the individual techniques has advantages and disadvantages, governments of several countries have already decided to use contactless RFID technology for future international passport documents. However, this decision does not prevent additional storage techniques to be implemented, thus the combination of various techniques from cryptography and data storage concepts as mentioned earlier appear to be feasible for providing additional security for ID document

6. POTENTIAL VULNERABILITIES AND COUNTER MEASURES

Potential security problems of technical systems can be differentiated with respect to the *ability*, the *possibility* and the *intention* of the attacker, the threatened security aspect or the attack techniques used.

On the one hand, threatened security aspects are the ones of the issuer of the ID documents (*document authenticity*, *non-transferability*, *integrity*) and on the other hand the ones of the document holder (*availability*, *confidentiality* as a *privacy* matter). An ID document is not authentic if the legitimate issuer is not the origin of the document. If someone else other than the legitimate holder presents an ID document as his own, the non-transferability is affected. An ID document loses its integrity if its content is modified or manipulated. Depending on the scenario and the ID document system an attacker could be able to distract the legitimate usage of the ID document, for example by sending jamming signals during contactless communications. The confidentiality of private data on a passport document could be violated, for example if an attacker is able to read them out.

With regards to their ability, possible attackers can be divided into corrupt *insiders* with special confidential knowledge, *outsiders* with high technical possibilities and knowledge and *others* without outstanding knowledge. These insiders could be employees of document issuing authorities, persons like border control officers, who have contact with technical installations for ID document inspection and other subjects with access to that insider knowledge, like the service personnel.

The possibilities to attack an ID document can be divided into *permanent*, *temporary*, *near-distance* and *mid-distance* access. Permanent access to an ID document implies the possibility to access invasively the stored data, for example by opening the used integrated circuit and manipulate it.[3, 24] Handling of a stolen passport is an example for permanent access. An example for temporary access to a document is a person handing out his passport and getting it back later, like for instance required at some hotel reception desks. Near-distance access as we understand it is for example the access to a RFID-based ID document using a reader/writer device within the standardized reading distance. By mid-distance access we mean for example the eavesdropping[12] of RFID to reader communication by enlarging the standardized reading distance. Figure 3 illustrates the three dimensions of possible spaces of vulnerability exploitation, the three axes denote the aspects of ability, possibility and security aspect and the subspaces denote potential attacks.

For example, a RFID Manipulation could be performed by outsiders, on mid to near distance and address the Integrity or document authenticity.

In this section we will analyze several potential vulnerabilities for one specific scenario: the expected integration of biometric features in European passport based on RFID technology[18], structured by the endangered security aspect.

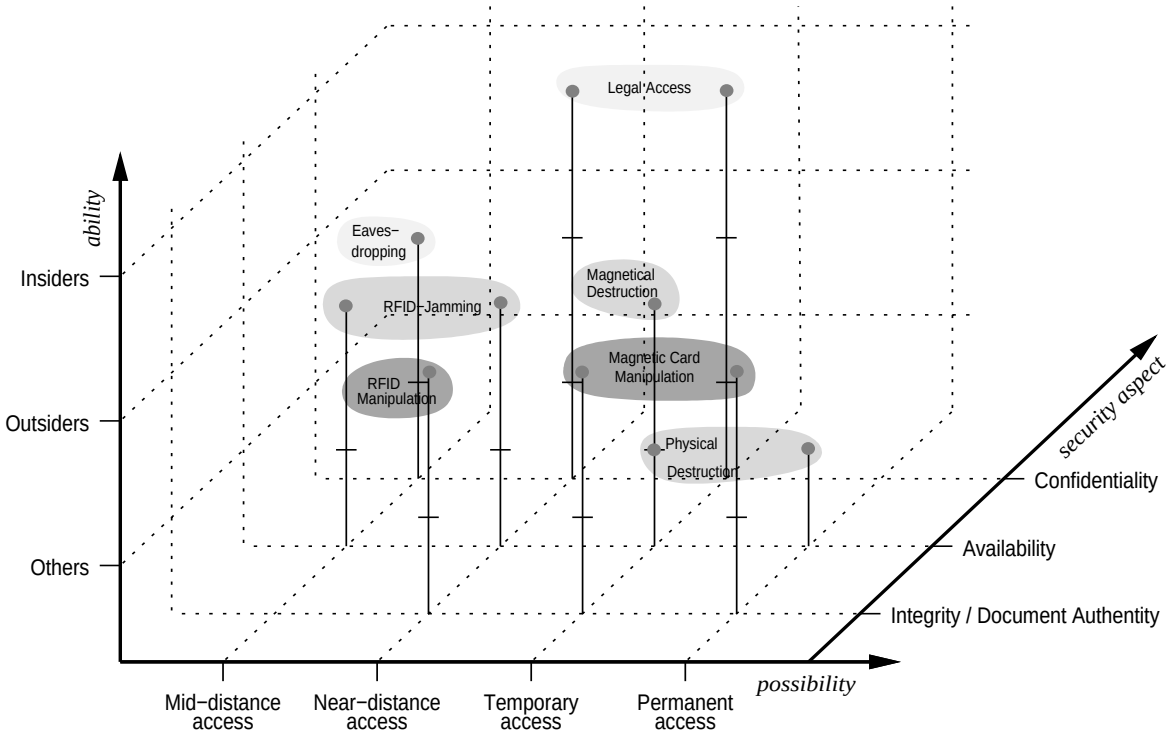

Figure 3: Dimensions of Vulnerabilities

6.1 Confidentiality

In the context of biometrics in ID documents, the security aspect of *confidentiality* is endangered mainly in two aspects. First, the saved information, be it biometric information like the face image, fingerprint or signature, or be it information like the name, date of birth and so on, could be considered as private data and be protected against unauthorized access.

The second privacy related issue in our context is the potential to create a profile of movement by tracking any kind of unique data, like the passport number of the holder*. Vulnerabilities in our first sense occur, if the information stored on the ID document is accessible to unauthorized personnel. One possible physical countermeasure to prevent breaches of confidentiality in context of ID documents can be to have an electric shielding in the shape of a metal cover for the passport document. This means that only if the holder of the document agrees the information can be read by some automated control mechanism. Only by willingly removing the metal cover, the data can be exchanged between the transponder and the reader.

Another possibility is to protect the biometric data by encryption, using some of the actually acquired biometric data as a password to open the biometric data stored in the passport. Of course, this requires the reliable construction of random keys from unreliable biometric data which is a nontrivial problem, however first promising approaches for key generation by voice, handwritten signature and fingerprint and have been presented.[30, 43, 44]

6.2 Availability

The availability of an ID document is endangered, if it is possible for a third person to prevent the legitimate holder of the document to use it in the intended way. This can be done by removing the digitally stored data, destroying the memory or jamming the communication between ID document and the reader device, either by electromagnetic means in case of contactless communication or by physical means. It is very much impossible to prevent an attack on the avail-

* *Consumer privacy may be compromised by nearby attackers extracting data from unprotected tags. Individuals may be physically detected by associating their identities with tags they carry, violating "location privacy" […]* [39]

ability by jamming the reader-transponder communication using high energy RF-radiation, although it appears quite unlikely that such proceeding would remain unperceived in observed environments such as border control stations. It should be relatively easy to track down such a source of energy and disable it eventually. To prevent an attack to the document using high-energy microwave radiation and therefore permanently destroying the document the above-mentioned metal shielding can be considered.

Besides physical protection, technical mechanisms can be used to ensure availability. Particularly for the communication between the RFID tag and the reading device, robust and error correcting communication protocols support compensation of noise or data package loss over the wireless transmission channel.

6.3 Integrity and Document Authenticity

To prevent attacks on the integrity and document authenticity such as replacing the stored biometric data by forged information (for instance by removing the original RFID-chip of a passport or making it permanently unusable by electrical or mechanical means and applying a forged one instead) an approach based on cryptographic combination of different storage methods, as described in section 5.10, could be used. It is generally not straightforward to forge information stored in print such as OCR-text and barcode, if the passport is protected using conventional physical measures such as special paper, laminating techniques, holograms and suchlike. Therefore, it appears possible to digitally sign the data stored electronically within the chip and embed that signature in the photograph with watermarking techniques or using OCR-text or barcodes to embed that information into the document and thus link the biometric data to other visible or readable information on it. However this protection scheme potentially allows for additional denial of service attacks, e.g. by removing the barcode containing the digital signature or tampering of the image in case of watermark storage.

6.4 Key Management

To enforce the data integrity and authenticity, as well as the confidentiality, the ICAO currently discusses using of public key cryptography for certified digital signatures and encryption of the stored passport data.[18]

However cryptographic solutions imply key management problems. In symmetric schemes, integrity and confidentiality can only be ensured to the degree that the keys can be considered secret. This is a hard constraint, considering the complexity of the international application of passports and the great number of persons involved in the processes of producing the ID documents, readers and backend systems. On the other hand, symmetric keys could be derived directly from machine-readable personal data on the document in this case the security would rely on the secrecy of the underlying algorithm, which does not seem acceptable as well. Using asymmetric cryptography may limit the problem, but implies a rather complex key management, where public keys of groups of users or even each individual user need to be managed and made available to all system participants. Again, considering the dimension of a large-scale implementation of biometrics in international travel documents, this does not appear viable.

6.5 Naive Approach

The naive approach for using biometrics in ID documents is to store the complete biometric data (e.g. the fingerprint iris, or face image) on a RFID transponder, embedded in the document. For verification, the biometric feature data T are acquired, the reference feature set R is read from the passport and T and R are compared and matched.

This approach bears a lot of problems; for the security of the system as well as for the holders' privacy. If there is no mechanism implemented to prevent manipulations of the biometric data on the transponder, everyone in the reading distance can remove or corrupt the data. Furthermore it is possible to save unauthorized new biometric reference features on the document, so that a third person would be able to fool a border control system. Even the passport holders' privacy is affected, if everyone in reading distance is potentially able to obtain private data. Having these biometric data, e.g. iris or fingerprint images, it could be possible to create artificial copies of that features, in order to present them to other biometric systems.

6.6 Biometric Hashes

To overcome some of the problems mentioned in section 6.5, an idea could be, not to store the whole biometric data but only an extract of them on an RFID transponder. In section 4, different approaches for extracting hash like data from biometric data are mentioned (e.g. BioHash for signatures or IrisCode). Saving only a kind of hash of the biometric feature data could prevent an attacker from forging the biometric feature. If furthermore the biometric reference features on the transponder have a digital signature of the authorized issuer of the passport, manipulations on the data can be detected. By additionally encrypting the stored data with a key, which is embedded in the passport using for example barcodes, OCR or watermarking in the face photograph, access to the biometric data is limited to persons with physical access to the document.

The problem of location privacy is not solved by these mechanisms. Also the used biometric hash function has to be reliably non-invertible, to prevent forgery of the biometric features from the hashed reference. The later problem could be met by using on-card matching techniques. The evaluation of these approaches will be part of our future work in this area.

7. CONCLUSIONS

In this article, we have reviewed the most important approaches to include digital biometric data in ID documents. We have provided an overview of the most adequate biometric features for this purpose and have shown that amongst these different modalities a trade-off problem between recognition accuracy, user acceptability and confidentiality exists. While the planned inclusion of face images in passport documents appears widely acceptable, it promises poor recognition accuracy, the usage of fingerprint may improve this aspect. However, there are privacy concerns for the latter approach and misuse and forgery do not seem unlikely. The inclusion of a third modality, the handwritten signature as active online feature seems viable as well.

With respect to the technical aspect of data storage techniques, we have shown that a number of different methods are available today and is seems reasonable for any ID card implementation to utilize combination of such techniques. Together with cryptographic functions such as hashes and digital signatures, this may ensure the security aspects of confidentiality, integrity and authenticity, whereas physical and signal coding techniques may be used to improve availability of the biometric information.

One possible solution to the problem, which we have referred to, is to derive cryptographic keys from each actually taken biometric sample, which is used to generate a reproducible individual key to protect the biometric reference. Although a few initial approaches have been referred to in this paper, this remains one of the future scientific challenges in biometrics. Also, the planned ubiquitous implementations in travel documents provide unique opportunities for future field evaluation of biometric algorithms and may result in statistically more significant statements regarding the capabilities of automated biometrics.

ACKNOWLEDGEMENTS

This work has been partly supported by the EU Network of Excellence SIMILAR (Proposal Reference Number: FP6−507609). The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the European Union.

REFERENCES

1. 3M-AiT Ltd., *Using 2D Barcodes to Enhance the Security of Machine-Readable Travel Documents (MRTDs)*, 2000.
2. A. Albrecht, U. Seidel, and M. Breitenstein, “Gesichtserkennung für den geplanten Einsatz in Lichtbildausweisen – BioP I”, in *Proceedings of D·A·C·H Security*, 2004.
3. R. Anderson, and M. Kuhn, “Tamper Resistance – a Cautionary Note”, in *The Second USENIX Workshop on Electronic Commerce Proceedings*, pp. 1–11, 1996.
4. BSI (Bundesamt für Sicherheit in der Informationstechnik), *BioFace – Comparative Study of Facial Recognition Systems*, *BioFace – Comparative Study of Facial Recognition Systems*, 2003.
5. BSI, *Untersuchung der Leistungsfähigkeit von Gesichtserkennungssystemen zum geplanten Einsatz in Lichtbilddokumenten – BioP I*, **http://www.bsi.bund.de/literat/studien/biop/biopabschluss.pdf** (in German), 2004.
6. Bundesdruckerei GmbH, *All security features of the passport data page*, **http://www.bundesdruckerei.de/en/iddok/2_1/2_1_7.html**, 2004.
7. R. Cappelli, A. Erol, D. Maio and D. Maltoni, “Synthetic Fingerprint-image Generation”, in *Proceedings of International Conference on Pattern Recognition (ICPR2000)*, Vol. 3, pp. 3475−3478, 2000.
8. I. Cox, M. Miller, J. Bloom, and M. Miller, *Digital Watermarking*, Morgan Kaufmann, ISBN 1-55860-714-5, 2001.
9. J. Daugman, “High confidence personal identification by rapid video analysis of iris texture”, in *Proceedings of the 1992 IEEE International Carnahan Conference on Security Technology*, pp. 1–11, 1992.
10. J. Daugman, “The importance of being random: Statistical principles of iris recognition”, in *Pattern Recognition*, Vol. 36, No. 2, pp. 279–291, 2003.
11. L. C. Ferri, A. Mayerhöfer, M. Frank, C. Vielhauer, and R. Steinmetz, “Biometric Authentication for ID Cards with Hologram Watermarks”, in *Proceedings of SPIE - Security and Watermarking of Multimedia Contents*, Vol. 4675, pp. 629−640, 2002.
12. T. Finke, and H. Kelter, *Abhörmöglichkeiten der Kommunikation zwischen Lesegerät und Transponder am Beispiel eines ISO14443-Systems*, **http://www.bsi.de/fachthem/rfid/Abh_RFID.pdf** (in German), 2004.
13. K. Finkenzeller, *RFID-Handbuch*, Carl Hanser Verlang, München, ISBN 3-4462-1278-7, 2002.
14. FVC2004: the Third International Fingerprint Verification Competition, **http://bias.csr.unibo.it/fvc2004**, requested December 2004.
15. Hitachi Ltd., *μ-Chip*, **http://www.hitachi.co.jp/Prod/mu-chip/**, 2003.

16. ICAO (International Civil Aviation Organization), *Technical Advisory Group on Machine Readable Travel Documents: Report of 14th Meeting*, **http://www.icao.int/mrtd/download/documents/TAG%2014%20-%20Report.pdf**, 2003.
17. ICAO, *Biometrics Deployment of Machine Readable Travel Documents – Technical Report*, Version 2.0 Final, 2004.
18. ICAO, *Technical Report – PKI for Machine Readable Travel Documents offering ICC read-only access*, Version 1.1, 2004.
19. ICAO, *Use of Contactless Integrated Circuits In Machine Readable Travel Documents – Annex I*, Version 4.0, 2004.
20. Infineon Technologies AG, *SLE 66CLX641P – 16-Bit High Security Contactless Controller*, 2004.
21. A. K. Jain, S. Prabhakar, and S. Pankanti, *Can Identical Twins be Discriminated Based on Fingerprints?*, 2002.
22. A. Jokers, *Hat der Fingerabdruck ausgedient?*, **http://www.heise.de/tp/deutsch/inhalt/lis/9711/1.html** (in German), 2001.
23. A. Juels, "Minimalist Cryptography for Low-Cost RFID Tags", in *Security of Communication Networks (SCN)*, 2004.
24. O. Kömerling, and M. G. Kuhn, "Design Principles for Tamper-Resistant Smartcard Processors", in *Proceedings of the USENIX Workshop on Smartcard Technology*, 1999.
25. M. G. Kuhn, and R. J. Anderson, "Soft Tempest: Hidden Data Transmission Using Electromagnetic Emanations", in *Information Hiding 1998*, pp. 124–142, 1998.
26. LaserCard Corp., **http://www.lasercard.com/**, 2004.
27. D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*, Springer-Verlag, New-York, 2003.
28. T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of Artificial "Gummy" Fingers on Fingerprint Systems", in *Proceedings of SPIE - Optical Security and Counterfeit Deterrence Techniques IV*, Vol. 4677, 2002.
29. T. Matsumoto, M. Hirabayashi, and K. Sato: "A Vulnerability Evaluation of Iris Matching (Part 3)", in *Proceedings of the 2004 Symposium on Cryptography and Information Security (SCIS)*, Institute of Electronics, Information and Communication Engineers, pp. 701–706, 2004.
30. F. Monrose, M.K. Reiter, Q. Li and S. Wetzel, "Using voice to generate crypto-graphic keys", in *2001: A Speaker Odyssey. The Speech Recognition Workshop*, 2001.
31. OPTEL, *Fingerprint Synthesis (Software)*, **http://www.optel.pl/software/english/synt.htm**, 2003.
32. Österreichische Staatsdruckerei (Austrian State Printing Office), *Security Features*, **http://www.staatsdruckerei.at/www/index_ie_en.html**, 2004.
33. T. Petermann, and A. Sauter, *Biometrische Identifikationssysteme (Biometric identification systems)*, TAB Working Report No. 76, **http://www.tab.fzk.de/de/projekt/zusammenfassung/ab76.pdf** (in German), 2002.
34. T. Petermann, C. Scherz, and A. Sauter, *Biometrie und Ausweisdokumente (Biometrics and identity documents)*, TAB Working Report No. 93, **http://www.tab.fzk.de/de/projekt/zusammenfassung/ab93.pdf** (in German), 2003.
35. Philips Electronics, P5CD072 – Secure Dual Interface PKI Smart Card Controller, 2004.
36. P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone: "Face Recognition Vendor Test 2002", in *NIST Technical Report*, **http://www.frvt.org/FRVT2002/documents.htm**, requested December 2004
37. J. Picard, C. Vielhauer, and N. Thorwirth, "Towards Fraud-Proof ID Documents using Multiple Data Hiding Technologies and Biometrics", in *SPIE Proceedings – Electronic Imaging, Security and Watermarking of Multimedia Contents VI*, pp. 123−234, 2004.
38. M. Sandström, *Liveness Detection in Fingerprint Recognition Systems*, M.Sc. Thesis at Linköping University, **http://www.ep.liu.se/exjobb/isy/2004/3557/exjobb.pdf**, 2004.
39. S. E. Sarma, S. A. Weis, and D. W. Engels, "Radio-Frequency Identification: Security Risks and Challenges", in *RSA Laboratories Cryptobytes*, Vol. 6, No. 1, 2003.
40. M. A. Sasse, "Assessing the Biometrics Enterprise: the present situation and future challenges", in *Proceedings of the IEE Seminar on the Challenge of Biometrics*, London, UK, ISBN 0-8634-1480-X, December 2004.
41. L. Thalheim, J. Krissler, and P. M. Ziegler, "Body Check − Biometric Access Protection Devices and their Programs Put to the Test", in *c't − Magazin für Computer und Technik*, Vol. 11, 2002.
42. A.S. Tolba, and A.N. Abu-Rezq, "Combined Classifiers for Invariant Face Recognition", in *Pattern Analysis Applications*, pp. 289–302, ISSN 1433-7541, 2001.
43. U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain, "Biometric Cryptosystems: Issues and Challenges", in *Proceedings of the IEEE, Special Issue on Enabling Security Technology for Digital Rights Management*, Vol. 92, No. 6, pp. 948–960, 2004.
44. C. Vielhauer, R. Steinmetz, and A. Mayerhöfer, "Biometric Hash based on Statistical Features of Online Signatures", in *Proceedings of the IEEE International Conference on Pattern Recognition (ICPR)*, Quebec City, Canada, Vol. 1, pp. 123–126, ISBN 0-7695-1696-3, 2002.
45. C. Vielhauer, *Handwriting Biometrics for User Authentication: Security Advances in Context of Digitizer Characteristics*, PhD Thesis, University Darmstadt, 2004.
46. D. Y. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, and G. Rigoll: "SVC2004: First International Signature Verification Competition", in *Proceedings of the International Conference on Biometric Authentication (ICBA)*, pp. 16–22, 2004.
47. *Legal Challenges to Fingerprints*, **http://onin.com/fp/daubert_links.html**, 2004.