Proceedings of the 2005 IEEE
Workshop on Information Assurance and Security
United States Military Academy, West Point, NY

Evaluation of Image Compression Algorithms for Fingerprint and Face Recognition Systems

Wolfgang Funk, Michael Arnold, Christoph Busch, Axel Munde

Abstract—

A variety of widely accepted and efficient compression methods do exist for still images. To name a few, there are standardised schemes like JPEG and JPEG2000 which are well suited for photorealistic true colour and grey scale images and usually operated in lossy mode to achieve high compression ratios. These schemes are well suited for images that are processed within face recognition systems.

In the case of forensic biometric systems, compression of fingerprint images has already been applied in Automatic Fingerprint Identification Systems (AFIS) applications, where the size of the digital fingerprint archives would be tremendous for uncompressed images. In these large scale applications Wavelet Scalar Qantization has a long tradition as an effective encoding scheme.

This paper gives an overview of the study *BioCompress* that has been conducted at Fraunhofer IGD on behalf of the Federal Office for Information Security (BSI).

Based on fingerprint and face image databases and different biometric algorithms we evaluated the impact of lossy compression algorithms on the recognition performance of biometric recognition systems.

I. Introduction

A variety of widely accepted and efficient compression methods do exist for still images. To name a few, there are standardised schemes like JPEG and JPEG2000 which are well suited for photorealistic true colour and grey scale images and usually operated in lossy mode to achieve high compression ratios. Others like LZW (e.g. used in the GIF image format) and RLE are designed for lossless compression that preserves the complete image information but provides only restricted compression capabilities. Lossless compression schemes are very useful for encoding binary and colour-indexed images.

In the case of biometric recognition systems, compression of fingerprint images has a long tradition in Automatic Fingerprint Identification Systems (AFIS) applications. These forensic systems interact with large scale archives with digital fingerprint images. As the sizes of these archives would be tremendous for uncompressed images, Wavelet Scalar Quantization (WSQ) has been established as a defacto standard for lossless yet high-quality compression of fingerprint images.

In contrast, non-AFIS identification systems such as physical access control systems to high-security areas or computing centers are generally operated with a limited number of enrolled subjects. They are associated with small databases and typically do not require compression of face or fingerprint images.

On the other hand, image size does matter for a verification scenario like the electronic passport (ePassport) application, where for practical and data privacy reasons the reference images are stored on a token such as a smart card. Many of these applications are open systems, where a major goal is independence of a specific recognition algorithm and additionally the option for visual inspection e.g. in false reject situation is required. In consequence storage of images inside the token is the prime choice.

This paper gives an overview of the study *BioCompress* that has been conducted at Fraunhofer IGD on behalf of the Federal Office for Information Security (BSI). The study investigated the impact of lossy compression algorithms on the performance of fingerprint and face recognition systems. In particular those compression algorithms that are referred to in the respective ISO/IEC JTC1 19794 documents were applied to a fingerprint and face image database, respectively. By varying the compression ratio the maximum acceptable level of compression for a particular combination of compression and recognition algorithm was determined.

This paper is structured as follows:

- Section II describes the ePassport application scenario and the scope of the study as well as related work.
- Section III introduces the design of the experiments conducted.
- Section IV describes the actual setup of the experiments.
- Section V reports the results of our experiments.
- Section VI summarizes the results of the study.

II. Scope and Background

The target scenario of the presented study was the storage of fingerprint and face images on personal tokens (e.g. smart cards) to be used by biometric recognition algorithms for verification purposes. There are several variations of this scheme:

W. Funk: Fraunhofer IGD, Darmstadt, Germany.
M. Arnold: Fraunhofer IGD, Darmstadt, Germany
C. Busch: Fraunhofer IGD, Darmstadt, Germany
A. Munde: Federal Office for Information Security (BSI), Bonn, Germany

0-7803-9290-6/05/$20.00 ©2005 IEEE.

- The image is retrieved from the token card and used as input to the template calculation (i.e. enrollment step) of a biometric recognition algorithm. The template is matched against the life image of the holder of the smart card.
- If the recognition algorithm is of low complexity and therefore suited to be stored and run on the token, a match-on-card application is possible.
- In the latter case it may be sufficient to store not the reference image of the biometric characteristic but the precomputed feature vector (template) on the card.

Each of the listed variations has its advantages and disadvantages. The first variation gives maximum flexibility regarding the algorithms to be used, as the algorithm is run outside the smart card. Moreover the image can be inspected by an expert which is especially relevant in false reject decision of the biometric system. The second variation also allows human inspection, but the number of algorithms that can be executed on todays smart card generation is limited. The same holds true for the third variation that uses templates and in most cases will not enable human inspection[1]

The obvious advantage of the match-on-card applications can be identified as increased data privacy, as the reference image or template does not leave the smart card. For the template-based approach very modest memory requirements for storing the user-specific features are a positive impact that must be balanced with an expected drop in the biometric performance. In that line the we can reflect the experiments of the International Labour Organization (ILO) The ILO had adopted a Convention in 2004 to improve the security for passengers and crews and the safety of ships. Implementing this Convention the ILO decided in early 2004 to augment Seafarers' Identity Documents (SID) with biometric characteristics. As the choice of the ILOs was a printed token with a two-dimensional PDF417 barcode the specification requires two fingerprints to be stored in a Minutia-Based template format [1]. As to the biometric performance the ILO evaluated various sensors and algorithms and reported that the systems under test could hardly meet the expected performance objectives and that the interoperability among vendors could basically not been reached [2].

In parallel the International Civil Aviation Organization (ICAO) in its New Technologies Working Group (NTWG) has analyzed biometrics for the purpose of machine assisted identiy confirmation of passport holders. The result of this analysis published was published in a series of technical reports [3],[4],[5], which will specify the ePassport as next generation of Machine Readable Travel Documents (MRTD)[6].

Some of the key findings of the reports are:

- To specify a biometrics technology for use at border control that would also allow visual inspection. In consequence facial images was choosen as primary identifier.
- To achieve a good biometric performance. In consequence facial images will be stored electronically in high capacity contactless IC media.
- To ensure high degree of interoperability and protect States against changing infrastructures or changing suppliers. In consequence reference samples will be stored as compressed images in the ePassport.

These findings were straigthforward at the point in time the ICAO completed the specification. However they introduced the ePassport trillemma. On the one side the preference for contactless access to the token resulted in the choice of IC to be conform to ISO 14443 (Radio Frequency Identification - RFID). These ICs allow access in the proximity up to 10 cm but they are limited in capacity: Current RFID can store a maximum of 72K. On the second side the choice for storage of reference images requires compression even in the case that just one facial image is stored (minimum requirement). For the member states of the European Union ePassports will include both facial images [7] and fingerprint images [8] On the third side of the trilemma, positive experience with face recognition in operational trials such as the Australian SmartGate System indicate that sufficient biometric performance can only be reached, if the intrinsic 2D face recognition pose problem is solved with multiple reference images from various viewpoints.

A similar situation holds true for fingerprint recognition systems. While feature extraction is generally invariant to rotation of the fingerprint, the impressions from the same source (subject) may stem from different parts of the finger. The acquisition of multiple impression covering the entire fingerprint is a simple workaround and generally solves the problem.

The request for multiple reference images strengthens the demand for high compression ratios and impacts back on the first side of the trilemma: Retrieving mulitple images from a RFID would linearliy increase the transaction time for the verification process that is at approx. 10s for one reference image.

However in the case that images are stored in the ePassport it is important to represent the image as compact as possible. Governments issuing ePassports will realized this by using state-of-the-art lossy image compression algorithms as specified in [8] and [7]. The most important task is to find the maximum compression ratio that does not pull the matching score below a predefined detection threshold.

As a consequence, the score value as a function of the compression ratio has to be determined. We selected 3 fingerprint recognition algorithms that were subject to 3 compression algorithms each, namely JPEG, JPEG2000 and WSQ. Moreover, we tested 2 face image recogni-

[1]In principle it is possible to reconstruct an image from the feature vector representing the template.

tion algorithms with 2 compression algorithms, JPEG and JPEG2000. We used a fingerprint database that was acquired at the Fraunhofer-IGD and a subset of the Feret face image database. Table I shows an overview of the test combinations.

	JPEG	JPEG2000	WSQ
Finger_1	x	x	x
Finger_2	x	x	x
Finger_3	x	x	x
Face_1	x	x	-
Face_2	x	x	-

TABLE I
TEST COMBINATIONS

JPEG is the most widely used compression method for photo-realistic images.
JPEG2000 is a successor to JPEG and provides better image quality than JPEG for very high compression ratios, at the cost of higher computational complexity.
WSQ is a standard for compression of fingerprint images that has been developed by NIST and is used to store the FBI fingerprint image database.

A similar study for face images has been conducted on behalf of Passports Australia [9], which has been extended by a detailed follow-up study [10]. This related work was based on "ICAO compliant" image material [3] where image sample were good representatives for facial images as they will be stored in future ePassport. On the contrary the image material in the Feret database can not be considered to be ICAO compliant. Therefore the focus of our study is on the detailed analysis of the impact of compression on fingerprint images and fingerprint recognition algorithms.

III. Design of Experiments

The tests are designed to find for each combination of recognition and compression algorithm the dependency of the score value on the compression rate and file size, respectively.

The idea behind the tests is to find the isolated impact of lossy compression on biometric recognition algorithms. The biometric features calculated from an image are matched with the features derived from a version of the same image that has been subject to lossy compression.

We evaluated two face recognition algorithms and three fingerprint algorithms.

Figure 1 shows the workflow each of the images is subject to at each of the compression rates under test. The image is compressed at a specific rate, decompressed and a template is calculated from the decoded image. The template is matched against the original image and the matching score value is recorded. The results for all images at a specific compression rate are used as input for one point of our measurement curve for a specific algorithm (cf. figure 2).

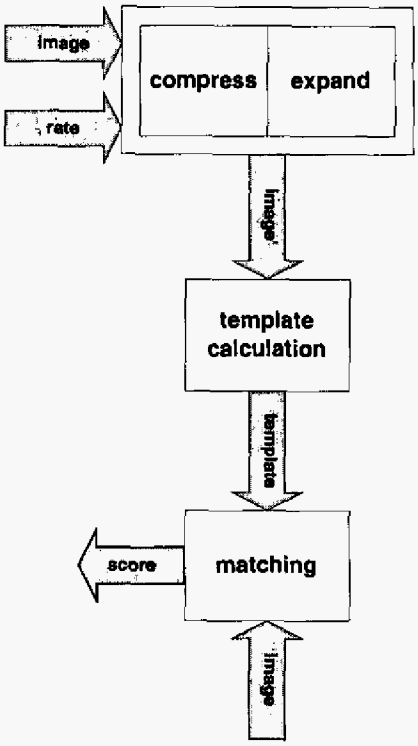


Fig. 1. Workflow for tests

We calculate the template from the compressed image. This is to make sure that if the enrollment module of the recognition algorithm applies additional processing steps prior to feature extraction, these are applied to the compressed version of the image. Moreover, this workflow matches the target application scenario, where the template will be calculated from the compressed image stored on a ePassport.

A. Metrics

This section briefly introduces the vocabulary used in the discussion of the test setup and the metrics used to analyse the results.

A.1 Matching Score and Threshold

The response of a matcher in a biometric recognition system is typically a *matching score* s (without loss of generality, ranging in the interval $[0, 1]$) that quantifies the similarity between the input and the database template representations. The closer the score is to 1, the more certain the system is that the two biometric features (e.g. fingerprint, facial image, image of iris, etc.) stem from the same source.

The system's decision is determined by the threshold T: Pairs of biometrical features generating scores higher than

or equal to T are classified as stemming from the same source (person).

A.2 Failure To Enroll (FTE)

This rate denotes the percentage of times users are not able to enroll in the recognition system. There is a trade-off between the FTE rate and the accuracy of a system. FTE errors typically occur when the recognition system performs a quality check to ensure that only good quality templates are stored in the database and rejects poor quality templates. As a result, the database contains only good quality templates and the system accuracy improves.

A.3 Metrics for Analysis

The mean of all score values for a particular compression ratio gives us one point on a plot which is expected to have a shape as shown in Figure 2. The mean score value plot shows at which compression ratios a considerable impact on the detection performance of the algorithms can be expected.

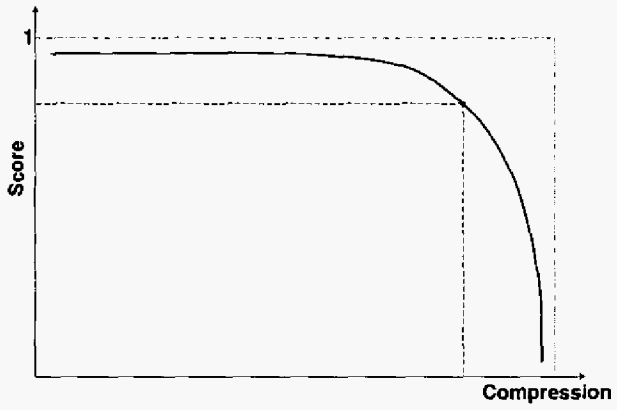


Fig. 2. Metrics for tests

The matching score is normalized so that in case of a perfect match in the uncompressed case the score value is exactly 1. The deviation from the perfect match value reveals the influence the compression has on the recognition algorithm. In general each uncompressed image will give a slightly different score value, thus we normalized on a per image basis. For each image and each particular compression ratio we divide the score value for the compressed case by the score value for the uncompressed case. The mean value of all normalized score values for a particular compression ratio gives us one point of the mean score value plot in figure 2.

IV. Setup

A. Fingerprint Database

The fingerprint database consists of 2160 images taken at a resolution of 500 dpi (dots per inch). The images show the complete fingertip on a light background and are 376 by 472 pixels in size. We acquired 3 fingerprint images from 30 individuals on 3 days each. Both hands and all fingers except the little finger were used. Thus we ended up with a database of 2160 fingerprint images. The 500 dpi sensor used for image acquisition was one of the best-performing sensors in the BioFinger I study and produced no enrollment errors (i.e. the FTE was 0) with the fingerprint recognition algorithms used in our tests. The images were stored in uncompressed TIFF format.

B. Face Database

We used a subset of the FERET face image database. The subset consists of 358 frontal uncompressed images in 24 bit colour with 512 by 768 pixels [11]. The images are stored uncompressed in PPM format.

C. Implementations of Compression Algorithms

We used the following software packages for our tests:

- The nconvert batch image conversion tool [12] which integrates the JPEG implementation of the Independent JPEG Group [13].
- Two different implementations of JPEG-2000 Part-1 [14], the JasPer software toolkit, Version 1.700.5 [15] and the LuraWave.jp2 command line tool, Version 1.03 [16].
- The NIST WSQ implementation for WSQ compression [17].

D. Actual Test Setup

D.1 Fingerprint Recognition

The tests were run on the complete fingerprint database, i.e. 2160 images. We selected 19 compression settings for JPEG, JPEG2000 and WSQ and each fingerprint recognition algorithm. Thus 2160 results per compression setting for a particular compression and recognition algorithm were obtained.

In contrast to JPEG2000, the JPEG and WSQ implementations (used in this study) can not predetermine the target file size exactly. Thus for JPEG and WSQ the mean file size values for each compression step were calculated.

D.2 Face Recognition

The tests were run on the subset of 358 frontal face images from the FERET database for 10 compression settings for JPEG and 11 settings for JPEG2000, respectively, and each face recognition algorithm. Thus we obtained 358 results for each combination of compression setting, face recognition algorithm and compression algorithm.

As for the fingerprint recognition case the JPEG file size was calculated as the mean file size for each compression setting.

V. Evaluation Results

The main experiment was to evaluate the impact of lossy compression on the detection score values that result from matching biometric images with biometric templates that have been created from compressed biometric images. To isolate the impact of compression, we matched templates derived from compressed images against the original images and calculated the mean value of the detection score for each compression ratio under consideration. This gave us the mean score value as function of the compression ratio and file size, respectively. The characteristics of this curve shows the impact lossy compression has on the detection capabilities of the recognition algorithm.

A. Fingerprint Recognition

The results for fingerprint recognition reveal that JPEG, JPEG2000 and WSQ perform equally well for our fingerprint database down to a file size well below 10,000 Bytes. Beginning at approximately 8,000 Bytes there is a considerable drop in recognition performance for JPEG. This behaviour is illustrated in the mean score value plots for the fingerprint recognition algorithms under test in Figure 3, 4 and 5.

Table II shows the normalized mean score values at a file size of 10,000 bytes, corresponding to compression ratio 0.056.

	JPEG	WSQ	JPEG2000
Finger_1	0.998	0.997	0.997
Finger_2	0.996	0.996	0.995
Finger_3	0.953	0.953	0.950

TABLE II
Normalized mean score values at file size 10,000 bytes

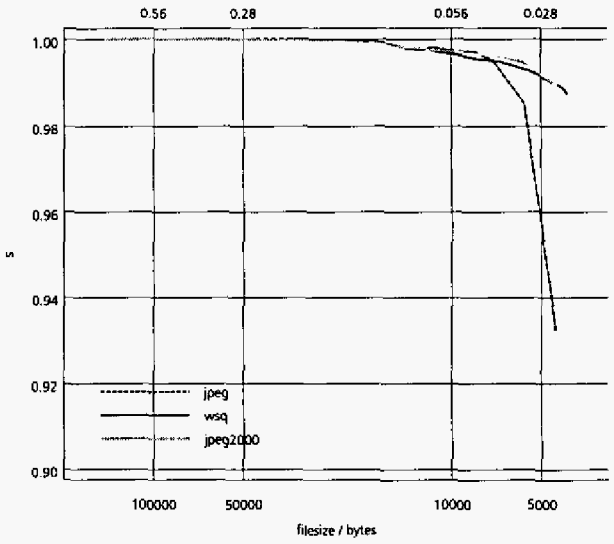


Fig. 3. Score values Finger_1

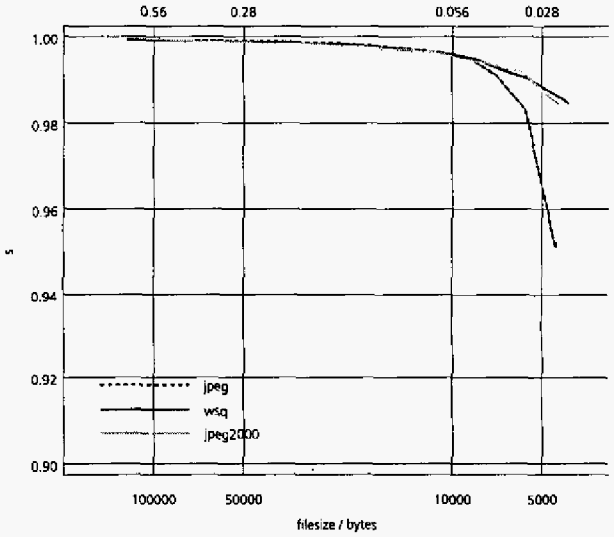


Fig. 4. Score values Finger_2

B. Face Recognition

As WSQ was not taken into account for the face recognition algorithms, we use a different plot style than in the previous section to compare the impact of JPEG and JPEG2000 on the score values. For each face recognition algorithm we show the relative JPEG performance: For each compression ration under consideration we divided the mean score value for JPEG by the mean score value for JPEG2000. Thus each plot shows the relative score value sr versus the file size and compression ration, respectively. For regions where the curve is beyond the line corresponding to $sr = 1$, JPEG performs better than JPEG2000 for the face recognition algorithm under consideration. Please note that for JPEG the target file size can not be exactly specified. Thus the abscissa values of the measuring points (i.e. the file size) for JPEG and JPEG2000 do not exactly match and the JPEG score values have been linearly interpolated at the JPEG2000 measuring points.

The results for face recognition presented in Figure 6 and Figure 7 show that JPEG performs similar to JPEG2000 for the Face_1 algorithm and even better than JPEG2000 for the Face_2 algorithm down to file size of 17,000 bytes [2], which is equivalent to a compression ratio of approximately 0.014. At higher compression ratios JPEG2000 seems more appropriate for the compression of facial images.

Moreover, it was interesting to see that for JPEG with

[2]It should be noted that this file size of 17K can not be compared directly with the ICAO studies [9], [10] as the size for uncompressed images is larger for samples in the Feret database.

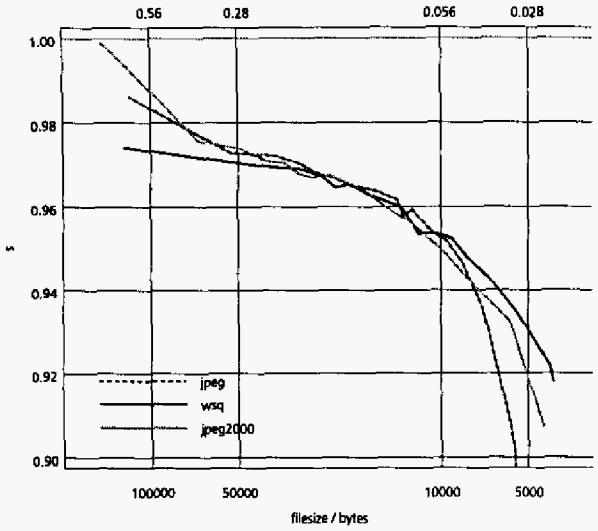


Fig. 5. Example score values Finger_3

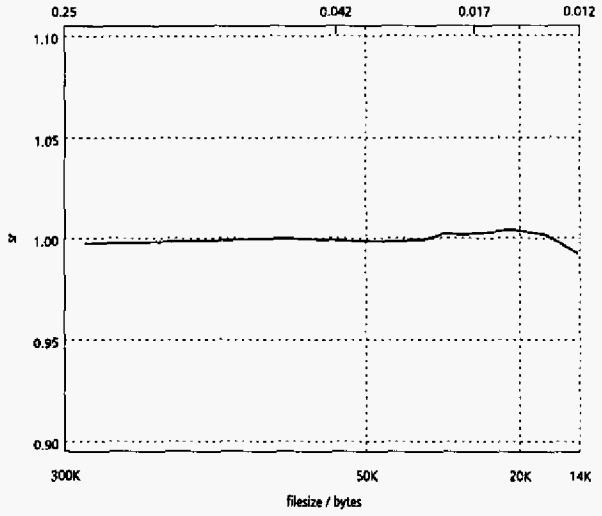


Fig. 6. Relative JPEG performance Face_1

high compression ratios either the score value decreases rapidly (Face_2) or a high number of failures to enroll does occur (Face_1). It seems that the Face_1 algorithm imposes stricter requirements on the image quality during the enrollment phase. Table III shows the absolute number of failures to enroll for Face_1 and JPEG versus the compressed file size in bytes for high compression ratios. It can be seen that for strong JPEG compression the majority of face images (303 out of 358 images) could not be used for enrollment. The Face_2 algorithm did not produce any failures to enroll.

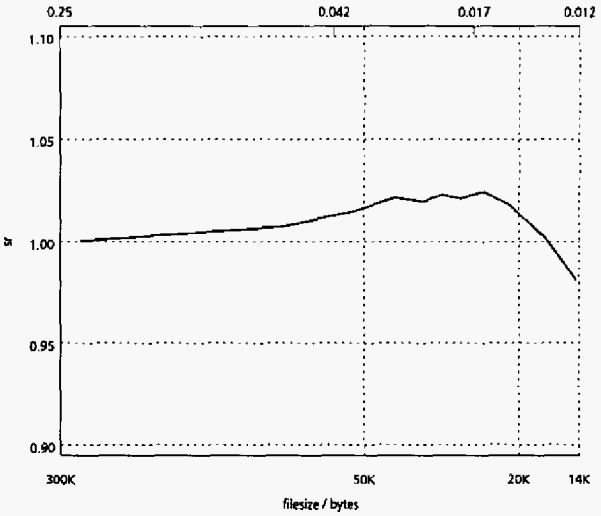


Fig. 7. Relative JPEG performance Face_2

Size	10700	15300	19600	23500	27500
Ratio	0.0091	0.013	0.017	0.020	0.023
Failures	303	18	1	1	2

TABLE III
FAILURES TO ENROLL FOR FACE_1 WITH JPEG COMPRESSION

C. Wavelet Filters

We ran tests with the fingerprint database to estimate the impact of different wavelet filter configurations for JPEG2000 on the recognition performance. JPEG2000 defines two different wavelet filters in Part 1 of the specification, the 5/3 and 9/7 filters. We used the Lurawave JPEG2000 implementation, which also provides a version of the 9/7 filter with adaptive quantization to evaluate the impact of different filters on the mean matching score.

Our results indicate that the choice of the wavelet filter has no significant impact on the detection performance. Even though the 9/7 filters showed a slightly higher performance at high compression ratios, the variations were clearly within the error bounds.

D. Region of Interest

Even though we did not test the Region of Interest (ROI) capabilities of the JPEG2000 libraries, it should be noted that this feature could significantly improve the compression ratio, once the background (outer region) is compressed at higher ratios then the interesting face (inner region) as defined in [7].

E. Error Resilience

To test the error resilience capabilities of JPEG2000 as compared to WSQ and JPEG, we randomly introduced

bit errors into a compressed fingerprint image. The modified image was decompressed and we compared the impact of the bit errors visually. Here a major advantage of JPEG2000 shows up. The impact of bit errors is considerably lower than for WSQ and especially for JPEG. Figure 8 in the appendix shows examples for a fingerprint image where 30 bit errors were introduced.

VI. Conclusions

We investigated the impact of lossy compression on the performance of fingerprint and face recognition algorithms. We evaluated two face recognition algorithms and three fingerprint algorithms, namely JPEG, JPEG2000 and WSQ.

The lossy compression algorithms that were investigated are already included in standardisation efforts for biometric data exchange formats, cf. the current ISO/IEC 19794 documents.

In our evaluation the impact of the JPEG algorithm on fingerprint recognition was comparable to the impact of JPEG2000 and WSQ, respectively down to a compression ratio of 0.056. For smaller ratios (i.e. for higher compression) WSQ and JPEG2000 are superior to JPEG.

The face recognition algorithms under test showed nearly the same behaviour for JPEG and JPEG2000 for a compression ratio down to 0.014. For smaller ratios, the JPEG2000 algorithm outperforms JPEG.

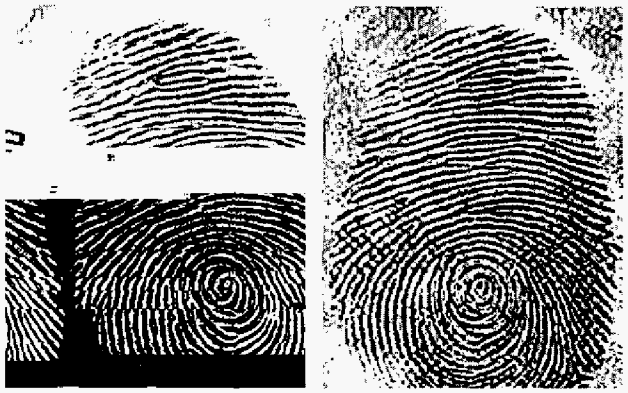
The tests of the fingerprint recognition algorithms were conducted with a fingerprint image database aquired with a 500 dpi sensor. Further research should consider using higher resolution fingerprint images, as now fingerprint sensor with a resolution of up to 1000 dpi are available.

References

[1] "Fdis 19794-2, biometric data interchange formats: Part2: Finger minutiae data," Jan. 2005.
[2] ILO, "Ilo seafarers' identity documents biometric testing campaign report - part i," tech. rep., International Labour Organization, nov 2004.
[3] ICAO, "Biometrics deployment of machine readable travel documents," tech. rep., International Civil Aviation Organization, May 2004.
[4] ICAO, "Contactless integrated circuit (ic) technical report," tech. rep., International Civil Aviation Organization, May 2004.
[5] ICAO, "Logical data structure technical report," tech. rep., International Civil Aviation Organization, May 2004.
[6] ICAO, "Machine readable travel documents," tech. rep., International Civil Aviation Organization.
[7] "Fdis 19794-5, biometric data interchange formats: Part5: Face image data," Jan. 2005.
[8] "Fdis 19794-4, biometric data interchange formats: Part4: Finger image data," Dec. 2004.
[9] ICAO, "Facial image optimal storage size - study 1," tech. rep., nov 2003. Biometrics Deployment Technical Report.
[10] ICAO, "Facial image optimal storage size - study 2," tech. rep., nov 2003. Biometrics Deployment Technical Report.
[11] NIST, "Feret - color feret, facial image database," tech. rep., Image Group, Information Access Division, ITL, National Institute of Standards and Technology, oct 2003. http://www.nist.gov/humanid/colorferet.
[12] P.-E. Gougelet, "NConvert image conversion tool," 2003. www.xnview.com.
[13] I. J. Group, "Independent JPEG Group." http://www.ijg.org/.
[14] International Organization for Standardization and International Electrotechnical Committee, *Information Technology - JPEG2000 - Image Coding System*, 2002.
[15] M. Adams, "The JasPer Project Home Page," 2003. http://www.ece.uvic.ca/~mdadams/jasper/.
[16] Algo Vision LuraTech GmbH, "LuraWave.jp2," 2004. http://www.algovision-luratech.com.
[17] R. M. M.D. Garris, C.I. Watson and C. Wilson, "Nfis - nist fingerprint image software," tech. rep., Image Group, National Institute of Standards and Technology, oct 2001.

VII. Appendix

A. Error Resilience Experiments



(a)JPEG with 30 bit errors (b)JPEG2000 with 30 bit errors

(c)Original

Fig. 8. Error Resilience