IET Biometrics

Research Article

IET Journals
The Institution of Engineering and Technology

Effects of image compression on ear biometrics

ISSN 2047-4938
Received on 23rd October 2015
Revised on 27th January 2016
Accepted on 15th February 2016
doi: 10.1049/iet-bmt.2015.0098
www.ietdl.org

Christian Rathgeb[1] ✉, *Anika Pflug*[2], *Johannes Wagner*[1], *Christoph Busch*[1]

[1]*da/sec – Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany*
[2]*Media Security and IT Forensics – Fraunhofer Institute for Secure Information Technology, Germany*
✉ *E-mail: christian.rathgeb@h-da.de*

Abstract: An ear recognition system represents a powerful tool in forensic applications. Even in case the facial characteristic of a suspect is partly or fully covered an image of the outer ear may suffice to reveal a subject's identity. In forensic scenarios imagery may stem from surveillance cameras of environments where image compression is common practice to overcome limitations of storage or transmission capacities. Yet, the impact of severe image compression on ear recognition has remained undocumented. In this work the authors analyse the influence of different state-of-the-art image compression standards on ear detection and ear recognition algorithms. Evaluations conducted on an uncompressed ear database are considered with respect to different stages in the processing chain of an ear recognition system where compression may be applied, representing the most relevant forensic scenarios. Experimental results are discussed in detail highlighting the potential and limitations of automated ear recognition in presence of image compression.

1 Introduction

The intricate structure of the outer ear constitutes a powerful biometric comprising highly discriminative biometric information. Generic automated 2D ear recognition systems consist of four major modules [1]: (i) image acquisition, where (in the ideal case) a side profile of a face image is captured; (ii) pre-processing, which involves ear detection and image enhancement. In past years, diverse approaches to ear detection have been proposed, for example, based on template matching [2, 3] or contour extraction [4, 5]; (iii) feature extraction, refers to the process in which the pre-processed ear image is reduced to a biometric feature vector summarising discriminative information. Numerous feature extraction algorithms have been suggested, for example, general purposes texture descriptors [6–8], intensity-based features [9, 10] or Gabor filters [11, 12]; (iv) comparison, where appropriate comparators, that is, classifiers, are used to obtain comparison scores. For surveys on ear recognition technologies the reader is referred to [1, 13]. In contrast to other well-established biometric characteristics, for example, fingerprints [14], iris [15], or face [16], the ear offers several advantages in the context of forensics. The morphology of the outer ear is fairly stable, and robust to change in facial expressions. Thus, ear recognition may outperform face recognition systems, especially in forensic scenarios where faces might intentionally be covered [1]. Profile images of subjects which comprise ears of sufficient quality can be extracted by the vast majority of today's surveillance cameras, as confirmed by a study on the evidential value of ear images from CCTV footage [17]. In addition, success stories in forensic investigations, for example, News [18], have confirmed the potential of ear recognition. The continuous recording of surveillance cameras generates large volumes of footage data that needs to be transferred from remote locations often via narrow bandwidth networks. Segmentation of data subjects and detectable biometric characteristics can reduce the data stream. Moreover, reduction of biometric data volume can be achieved by compression of captured images. However, in order to provide vendor-neutrality the International Organisation for Standardisation (ISO) specifies different biometric characteristics to be recorded and stored in (raw) image format rather than in extracted reference data (templates). With respect to standardisation ISO/IEC IS 19794 on 'Biometric Data Interchange Formats' [19] represents the most relevant standard. Based on several empirical evaluations on the influence of image compression on fingerprints, for example, in [20–22], face, for example, in [23–26] or iris, for example, in [27–30], the JPEG 2000 image compression standard [31] is recommended for compression of face and iris images. While conventional fingerprint recognition systems store fingerprints in a minutiae-based format forensic applications require fingerprint images acquired at 500 to 1000 ppi to be compressed with JPEG 2000 [19, 32]. In the vast majority of studies it is found that, depending on biometric characteristics, fairly high compression of biometric data can be achieved, while biometric performance is maintained [22, 26, 33]. Further, it has been observed that lossy compression causes effects similar to low-pass filtering which may even improve recognition accuracy [26, 28]. It is important to note that appropriate bit-rates may depend on the image resolution since low-resolution images may not allow for further compression. Further, obtained bit-rates depend on whether some kind of pre-processing is applied, for example, higher compression may be applied to original face images while compression at identical bit-rates may cause a significant decrease in performance when applied to cropped regions of interest (RoIs) since homogeneous regions in original images may allow for higher compression [26]. Similar effects have been confirmed for iris where higher compression can be employed on cropped images of an eye compared with pre-processed (polar) iris images [33]. Such observations clearly hamper an explicit comparison of obtained results reported for (different) biometric characteristics.

Despite investigations on compression of raw fingerprints images which are supported by investigation authorities, for example, the Federal Bureau of Investigation [34], existing studies on the effects of image compression on biometric data may be of less relevance from a forensic perspective. Aforementioned studies as well as standardisation activities analyse the effects of different image compression standards on biometric recognition systems in order to elaborate a common standard for the compression of biometric data which still enables reliable authentication [19]. Further, diverse compression techniques have been explicitly designed for distinct biometric modalities, for example, in [20, 35–38]. Hence, in various application scenarios lossless or slight compression is considered in bi-lateral compression scenarios, that is, probe and reference images

are compressed. In contrast, forensic applications may also be presented with highly degraded images caused by lossy compression. Moreover, comparisons of bad-quality (compressed) probe images to good-quality (uncompressed) reference images, that is, uni-lateral compression scenarios, may also represent a likely case in forensic investigations. Further, for forensic investigations the movement of unknown data subjects (not registered in a database) between different camera positions needs to be tracked then comparison of highly compressed samples stemming from different footage material must be analysed, that is, bi-lateral compression scenarios are of relevance as well.

Focusing on ear biometrics, so far, no studies investigating the effects of image compression have been proposed. In [39, 40] the impact of noise and blur on ear detection and recognition algorithms has been investigated. In these works experiments simulated camera noise and out-of-focus blur. It was found that while conventional detection algorithms appear rather robust in presence of these signal degradations, recognition algorithms are highly influenced by noise causing a significant decrease in accuracy for images exhibiting peak signal to noise ratio (PSNR) values of 20–30 dB. Further, effects of barrel distortion and super-resolution on ear biometrics are investigated in [41]. It has been empirically shown that biometric performance drops in case of barrel distorted images, which might be compensated by the use of super-resolution images for model normal vector creation. While these studies merely focus on effects caused by potential image distortions, omnipresent artefacts caused by image compression represent a more relevant signal degradation which further motivates our work.

In this work we study the effects of image compression on ear biometric recognition systems. We mainly consider two widely used image compression standards, JPEG [42] (JPG) and JPEG 2000 [31] (J2K). In addition, we present experiments on alternative compression standards, that is, JPEG XR [43] (JXR) and the recently proposed better portable graphics [44] (BPG). Uncompressed samples of different widely used ear databases are compressed at various bit-rates and the influence of considered image compression standards is reported for a scenario-based evaluation. Thereby, simulated recognition scenarios in the presence of severe image compression are comparable with corresponding original (un-effected) systems. In contrast to existing studies on the impact of image compression on biometric recognition emphasis is put on scenarios that are considered as most relevant for forensic applications. Depending on the scenario image compression is applied at different stages of an ear biometric processing chain. In order to cover all considered uni- and bi-lateral scenarios we provide a detailed investigation of the impact of image compression on different state-of-the-art ear detection algorithms as well as feature extractors. Hence, for experts in the field of forensics our evaluations provide a comprehensive reference on the feasibility and limitations of (semi-)automated ear recognition under the influence of image compression.

The remainder of this article is organised as follows: Section 2 summarises the considered image compression standards, the employed ear database, and the experimental protocol. In Section 3 different scenarios are introduced, evaluated, and discussed in detail. In Section 4 further investigations considering different image scales and alternative compression techniques are presented. Conclusions are drawn in Section 5.

2 Image compression and experimental setup

In the following sections we describe the employed image compression standards and specific compression artefacts frequently caused by them. Subsequently, we introduce the employed merged dataset, according ground truth information for ear detection, and the performance metrics used in experimental evaluations. Further, we analyse the quality of compressed images in terms of rate distortion performance.

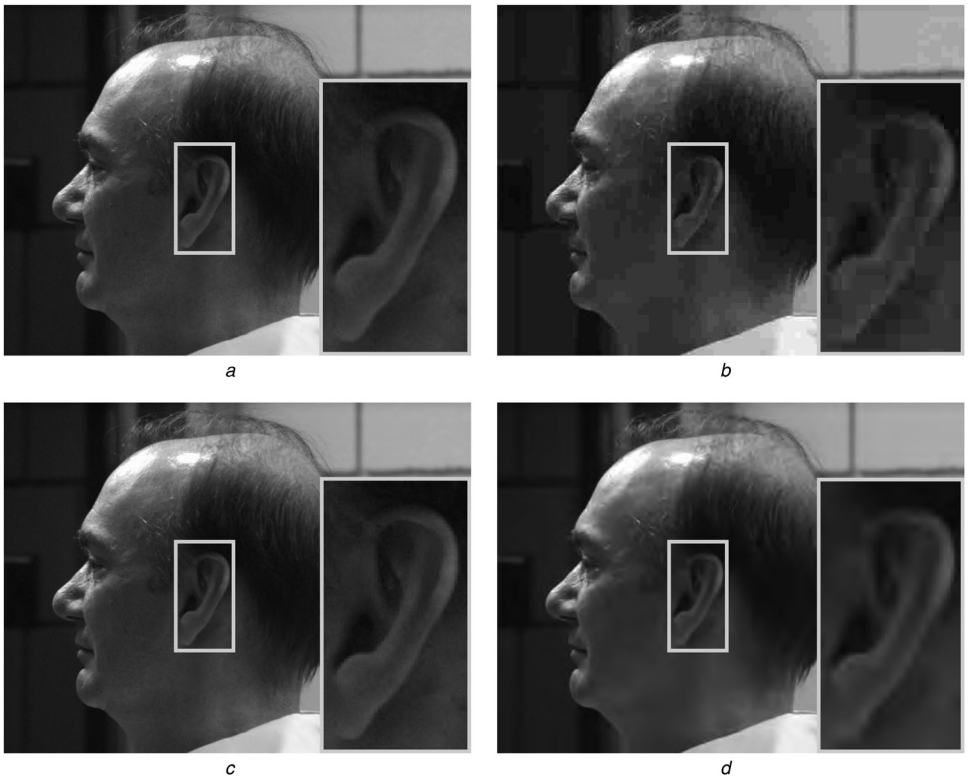

Fig. 1 *Sample compressions of image ID 02463d002 of the UND-J2 ear database [47] and close-ups of RoIs (at the right-hand side bottom of each subfigure) at compression rates of 1 and 0.1 bpp. Homogeneous regions within the original image (e.g. at the right-hand side top) allow for a more efficient image compression*

a JPG 1 bpp
b JPG 0.1 bpp
c J2K 1 bpp
a J2K 0.1 bpp

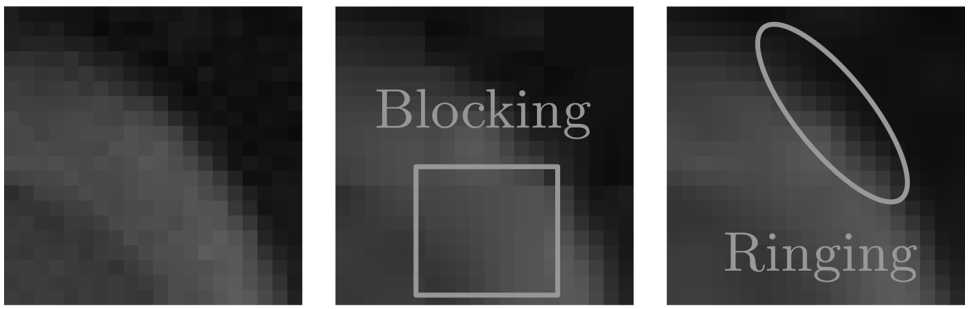

Fig. 2 *Comparison of artefacts: 20 × 20 pixel patches at the helix of the ear of image ID 02463d002 at a compression rate of 0.2 bpp*

a Original
b JPG
c J2K

2.1 Image compression standards and artefacts

As previously mentioned, in the majority of experiments we employ two lossy image compression algorithms (in default settings) for bitrates between 0.1 bits per pixel (bpp) and 1.0 bpp:

- *JPEG (JPG)*: the widely-used DCT-based image compression method is defined in ISO/IEC IS 10918-1 [42]. We iteratively configure quality parameters in order to obtain desired bit-rates using the ImageMagick software *convert* [45].
- *JPEG 2000 (J2K)*: the wavelet-based image compression standard which operates at higher compression ratios is defined in ISO/IEC IS 15444-1 [31]. The *JJ 2000* software [46] is employed as conversion tool which allows for explicit rate control.

Compressions for the minimum and maximum bit-rates for a sample image of the employed database (see Section 2.2) are shown in Fig. 1. Fig. 2 illustrates an example of artefacts resulting from JPG and J2K compression. When performing block-based coding for quantisation, as in JPG-compressed images, ringing and block boundary artefacts can appear. Results of J2K compression typically do not generate block-based artefacts of the original DCT-based JPG standard. J2K only produces ringing artefacts, manifested as blur and rings near edges. Another difference between JPG and J2K is that the latter requires complex and computationally more demanding encoders/decoders [31].

MPEG or Motion JPEG2000 (MJ2) encoded video footage comprises identical artefacts, respectively. In this work we only consider still images while, in addition, other compression artefacts may occur in video footage such as block boundary discontinuities at edges of motion compensation prediction blocks even in case error-protection is available, for example, in-loop de-blocking filter in H.264/AVC [48].

2.2 Databases, ground truth and experimental protocol

For our evaluation, we have composed a dataset of mutually different images of the publicly available UND-J2 [47] and UND-NDOff-2007 [49] ear database. The merged dataset contains a total number of 2192 left profile images of size 640 × 480 pixels from 513 subjects with least 5 and up to 17 images per subject. In contrast to existing works, in which experiments are conducted on significantly smaller datasets, for example, <100 subjects, this fused dataset allows for a more comprehensive and hence statistically more relevant evaluation. Only yaw poses of -60° and -90° degrees are considered as relevant, since off-angle images, which might not pass quality checks within automated ear recognition or forensic investigations, are expected to cause a significant drop in recognition accuracy [50]. Performance is evaluated by conducting a total number of 512 genuine and 512 × 513 = 262,656 impostor comparisons per cross-validation. For a single comparison four reference images per subject are randomly chosen and compared against the given probe image, where the median score is returned. We perform a 10-fold cross-validation based on which the median results are reported. Thereby, sampling errors are minimised while statistical significance increases. For the entire database ears were manually annotated in form of ear bounding boxes in order to obtain an appropriate ground truth for automated ear detection algorithms. The annotated bounding boxes yield an average size of 125 × 95 pixels for the entire data set.

Table 1 Examples of state-of-the-art camera models

Vendor	Product	Focal length, mm	Resolution	Sensor
ACTi[a]	D82	2.8–12	1920 × 1080	1/3.2″
AXIS[b]	P3367V	3–9	1920 × 1080	1/3.2″
GeoVision[c]	GV-FD220G	3–9	1920 × 1080	1/2.5″
Veilux[d]	VVIP-2L2812	2.8–12	1920 × 1080	1/2.5″

[a]http://www.acti.com/
[b]http://www.axis.com/
[c]http://www.geovision.com.tw/
[d]http://www.veilux.net/

In [51] the average size of the outer ear of males and females across different age groups is measured as 61.7 × 37.0 mm and 57.8 × 34.5 mm, respectively. For an average angle of auricle of 32.5° across age groups and sex we approximate the bounding box of an ear of any subject as 70 × 60 mm. Examples of state-of-the-art surveillance camera models made available by major vendors are shown in Table 1 (characteristics refer to currently best products). One could assume a camera with focal length of 8 mm, sensor diagonal of 1/2.5″, and resolution 1920 × 1080 pixels. The diagonal D of the field of view at a distinct distance is estimated as, $D = A \cdot \tan(2 \cdot \arctan((d/2f)/2)) = A \cdot d/2f$ for focal length f and sensor diagonal d. For the considered camera model the aspect ratio is 16:9, that is, the field of view in object space corresponds to $16 \cdot \sqrt{D^2/(16^2 + 9^2)}m \times 9 \cdot \sqrt{D^2/(16^2 + 9^2)}$. This means, the considered camera model would capture images where ear regions comprise approximately 125 × 95 pixels at a distance of $A \simeq 2$ m, the same procedure has been applied in [39, 40]. We consider ∼2 m as reasonable distance for ear recognition while smaller re-scaled images are considered in further experiments, see Section 4.1.

Performance is estimated in terms of equal error rate (EER) and (true-positive) identification rate (IR), in accordance to the ISO/IEC IS 19795-1 [52]. In experiments identification is performed in the closed-set scenario returning the rank-1 candidate as identified subject (without applying a decision threshold).

2.3 Image quality estimation

To quantify the quality of compressed images we apply PSNR and structural similarity index measure (SSIM) by Wang *et al.* [53]. PSNR is still widely used because it is unrivalled in speed and ease of use while SSIM also takes structural features into account. Fig. 3 depicts estimated average quality for JPG and J2K compression of original images as well as according RoI, that is, manually annotated bounding boxes comprising outer ears (cf.

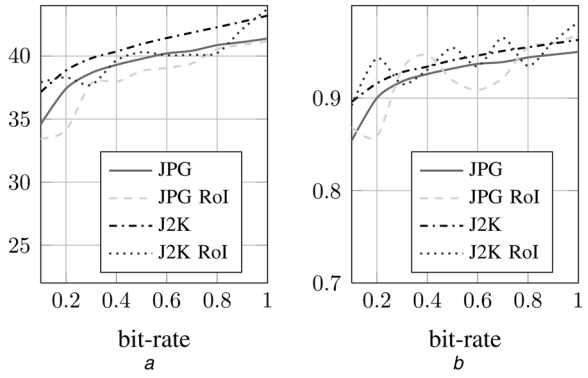

Fig. 3 *Quality estimation in terms of average PSNR and SSIM for original compressed images and according RoIs*

a PSNR (dB)
b SSIM (score)

Fig. 1), of the employed database. When estimating PSNR and SSIM values for the entire image a clear trend across bit-rates is observed. Typical values for the PSNR in lossy image and video compression are between 30 and 50 dB and acceptable values for wireless transmission quality loss are considered to be about 20 to 25 dB [54], hence, obtained image quality values for original images appear fairly good. Further, it can be observed that J2K consistently outperforms JPG with respect to rate-distortion performance. It is important to note that for original images more efficient compression may be achieved due to homogeneous image regions. Focusing on RoIs only, PSNR values tend to be lower and PSNR as well as SSIM values vary more for different bit-rates. This means, for distinct bit-rates cause a gain of rate-distortion performance within RoIs. These compression rates might have a beneficial effect on biometric performance. However, rate-distortion performance may not serve as adequate predictor for biometric performance [40, 55].

3 Evaluation scenarios

We consider four different scenarios including uni- and bi-lateral compression where ear detection is either performed automatically or by using ground-truth data. A flow-diagram of all scenarios is shown in Fig. 4. As can be seen, image compression is applied to original images prior to segmentation in all scenarios. The following sections describe applied feature extraction techniques, detection algorithms and each compression scenario in detail.

3.1 Feature extraction

In our experiments we apply four well-established statistical image feature extractors. Each feature extractor processes pixel-neighbourhoods within image patches resulting in a sequence of histogram distributions (one per patch) forming the biometric feature vector:

- *Local binary patterns (LBP) [56, 57]:* We extract LBP features from the n-8 neighbourhood of each pixel in the image. We divide the image into a regular grid of 10×10 pixels and concatenate the local LBP histogram within each grid cell.
- *Local phase quantisation (LPQ) [58, 59]:* We extract local LPQ histograms from a regular grid with 11×11 pixels cells and concatenate each of the local histograms to obtain the global feature vector.
- *Histogram of oriented gradients (HOG) [60]:* The HOG descriptor in our experiment is using a patch size of 8×8 pixels with nine different orientations.
- *Binarised statistical image features (BSIF) [61]:* Features are extracted using 5×5 pixels filters of 5-bit depth on 20×20 pixels sub-windows.

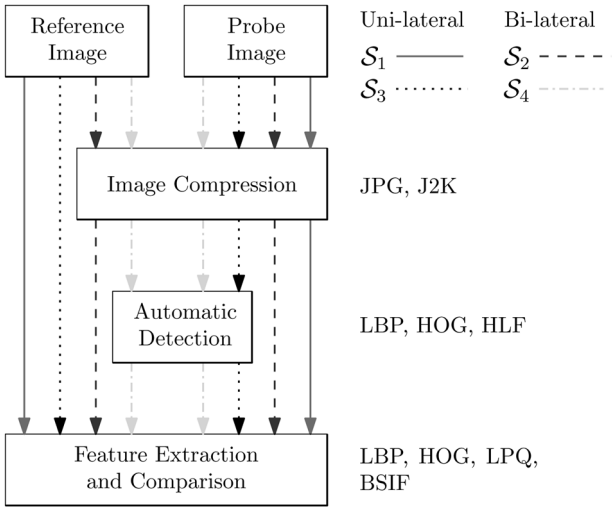

Fig. 4 *Overview of four considered scenarios ($\mathcal{S}_1$–$\mathcal{S}_4$) and according application of image compression and automated segmentation of reference and probe images*

We have chosen spatial histogram features since these directly encode distinctive properties of the image signal, for example, gradients or phase, without any further semantic interpretation. More recently, it has been shown that general purpose texture descriptor, for example, LPQ or BSIF, which process images in a patch-wise manner, maintain or even improve biometric performance compared state-of-the-art methods, for example, Gabor filters, in real-world conditions where variations in illumination and pose frequently occur [8]. Similar observations have been made for other biometric characteristics, for example, palmprint [62]. High-level features, such as outlines or landmarks are usually based on the same low-level image information as spatial histograms. Consequentially, we assume that effects of image compression on feature descriptors based on spatial histograms relate to changes in high-level features, for example, inaccurately positioned landmarks or wrongly detected outlines. Employed appearance features exhibit a large number of dimensions. For creating our feature space, we compute a projection matrix based on our training data by using LDA. After computing the feature space from the training images, we project the test images into the same feature space. Subsequently, we assign an identity based on a NN-classifier and cosine distance. The source code for feature space projection and classification is based on the PhD face recognition toolbox [63].

3.2 Scenario 1: compressed probe – manually segmented

In the first uni-lateral scenario ($\mathcal{S}_1$) the ear recognition system is presented with a compressed probe image where a database of good-quality (uncompressed) correctly segmented reference images is available for comparison. In addition the outer ear within the probe image is manually segmented, that is, the ground truth data is employed for both, probe and reference images. $\mathcal{S}_1$ refers to the likely case where the subject to be identified has already been found in acquired image material, for example, in case of a robbery where video footage is captured directly by surveillance installed at the crime scene. We consider $\mathcal{S}_1$ as the most relevant scenario from a forensic perspective.

Obtained performance rates for verification and identification for $\mathcal{S}_1$ are summarised in Table 2. For this scenario identification performance is considered as more relevant. The relative increase or decrease in terms of EERs and IRs for the best performing feature extractors is shown in Fig. 5. Focusing on the baseline performance (uncomp.) we observe that LPQ and BSIF descriptors yield best results followed by HOG and LBP. In general recognition accuracy tends to decrease in this uni-lateral compression scenario, however,

Table 2 Performance (EER/IR) of LBP, HOG, LPQ, BSIF for JPG, J2K (Scenario 1)

Format	bpp	LBP	HOG	LPQ	BSIF
uncomp.	–	7.03/74.01	8.97/74.77	4.36/82.26	5.05/79.20
JPG	1.0	8.38/69.11	9.06/74.77	4.75/82.11	5.06/77.52
	0.9	8.39/69.11	9.06/74.77	4.75/82.11	5.06/77.52
	0.8	8.98/64.98	9.19/74.01	4.50/82.42	5.54/79.36
	0.7	9.16/64.22	9.68/72.94	4.70/82.11	5.62/77.98
	0.6	9.60/61.77	9.48/71.25	4.71/81.80	6.23/75.84
	0.5	11.04/58.27	9.41/70.18	4.62/82.72	5.96/75.23
	0.4	12.36/54.13	10.13/67.74	4.44/81.19	6.73/73.55
	0.3	13.39/44.04	11.03/65.29	4.90/81.96	7.27/70.49
	0.2	17.17/32.72	13.02/57.03	5.05/81.50	9.74/60.24
	0.1	26.01/17.43	21.76/26.30	6.87/73.55	18.26/35.63
J2K	1.0	8.34/69.11	9.14/73.24	4.70/81.65	5.75/77.83
	0.9	8.78/67.58	9.30/72.48	4.62/81.50	5.66/77.52
	0.8	8.77/63.76	9.54/72.78	4.59/81.50	5.77/77.52
	0.7	9.57/61.31	9.25/72.94	4.65/81.96	5.92/76.61
	0.6	10.37/58.10	9.49/70.64	4.66/82.11	6.28/75.84
	0.5	11.25/54.28	9.43/69.72	4.63/81.65	6.88/73.70
	0.4	12.44/52.14	10.08/69.57	4.63/82.11	7.34/70.95
	0.3	13.46/48.78	10.27/68.04	4.92/80.73	8.28/64.68
	0.2	15.26/42.20	11.02/64.98	4.90/80.12	9.28/57.19
	0.1	18.04/35.02	13.26/54.59	5.56/79.05	12.09/44.34

Table 3 Performance (EER/IR) of LBP, HOG, LPQ, BSIF for JPG, J2K (Scenario 2)

Format	bpp	LBP	HOG	LPQ	BSIF
uncomp.	–	7.03/74.01	8.97/74.77	4.36/82.26	5.05/79.20
JPG	1.0	6.45/75.84	8.21/75.23	4.35/82.72	5.39/77.83
	0.9	6.45/75.84	8.21/75.23	4.35/82.72	5.39/77.83
	0.8	6.87/76.61	8.87/74.31	4.54/82.57	5.58/78.59
	0.7	7.25/75.69	9.32/74.31	4.51/82.11	6.15/78.13
	0.6	6.86/73.09	8.69/73.85	4.85/81.80	6.00/76.45
	0.5	7.22/71.25	9.13/74.16	4.68/81.04	5.96/78.75
	0.4	7.28/72.78	9.36/73.55	4.55/81.04	6.27/75.38
	0.3	9.24/69.11	9.31/70.95	5.08/80.12	6.68/74.46
	0.2	10.34/62.54	10.11/66.67	5.59/76.30	8.27/67.58
	0.1	16.66/44.04	15.63/44.03	10.30/48.93	13.61/45.26
J2K	1.0	7.09/74.01	8.83/74.31	4.74/82.57	5.39/78.75
	0.9	6.81/75.69	8.95/74.01	4.71/82.72	5.17/78.75
	0.8	6.46/73.85	8.89/75.38	4.82/82.57	5.20/78.75
	0.7	6.85/75.38	8.82/74.46	4.32/81.80	5.16/79.66
	0.6	6.95/75.38	9.22/75.54	4.94/81.19	5.01/79.97
	0.5	6.36/76.91	8.57/73.85	4.97/81.04	5.49/79.97
	0.4	7.54/74.92	9.05/74.31	4.97/81.35	5.76/78.59
	0.3	8.15/74.01	8.77/74.92	5.11/79.97	5.77/76.61
	0.2	7.68/69.72	9.36/72.17	5.75/77.22	6.15/74.46
	0.1	9.91/64.37	9.75/66.97	7.19/72.17	8.62/67.58

all feature extraction techniques appear rather robust down to bit-rates of 0.3 bpp while the performance of HOG and LBP descriptors is affected more compared with LPQ and BSIF. Further, we observe that for reasonable image quality, that is, down to 0.3 bpp (cf. Fig. 3), comparable biometric performance is obtained for JPG and J2K compression. That is, for the biometric performance of an ear recognition system the choice of feature extraction appears to be more important than the choice of image compression algorithm, that is, feature descriptors such as LPQ or BSIF which are designed to extract robust phase information generally reveal better performance compared with techniques employing merely gradient information such as HOG or LBP.

3.3 Scenario 2: compressed probe and reference – manually segmented

In the second scenario ($\mathcal{S}_2$) we analyse a bi-lateral compression, that is, probe as well as reference images are compressed at the same bit-rate. Again, segmentation is performed based on the available ground truth information. That is, $\mathcal{S}_2$ refers to the case where pairs of ear images to be compared are extracted from similar sources, for example, different surveillance cameras where the outer ears of subjects are manually segmented. Such a scenario may be relevant in forensic investigations where the presence of subjects at different places is confirmed without or prior to verifying their identities, for example, confirming that a single subject appears at different crime scenes (tracking).

Table 3 summarises performance rates for $\mathcal{S}_2$. In contrast to $\mathcal{S}_1$, in $\mathcal{S}_2$ verification performance may be more important. Relative changes in EERs and IRs for the two best performing feature extractors are depicted in Fig. 6. First, we observe that the ranking of feature extraction algorithms is the same as in $\mathcal{S}_1$. However, in this bi-lateral scenario biometric performance tends to improve for the majority feature descriptors across considered bit-rates, in particular for HOG and LBP. Similar effects have already been observed for other biometric characteristics, for example, for iris [33]. On the one hand, de-noising capabilities of slight image compression can have

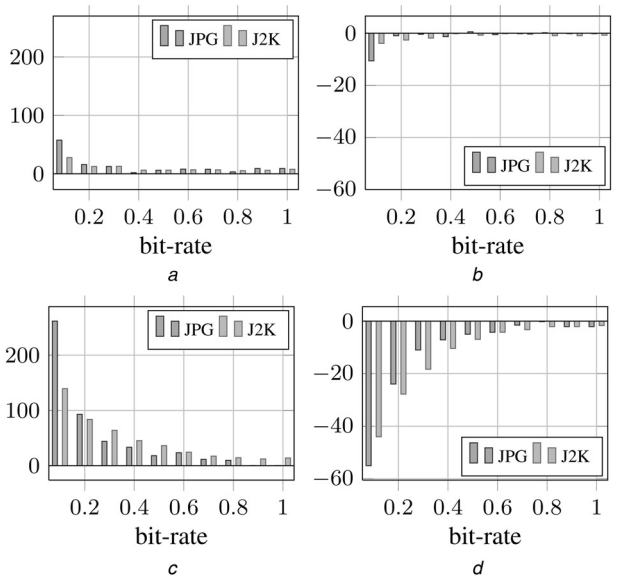

Fig. 5 *Relative change in EERs and IRs compared with the baseline performance for $\mathcal{S}_1$ at different bit rates for LPQ and BSIF*

a ⇅ EER (%) LPQ
b ⇅ IR (%) LPQ
c ⇅ EER (%) BSIF
d ⇅ IR (%) BSIF

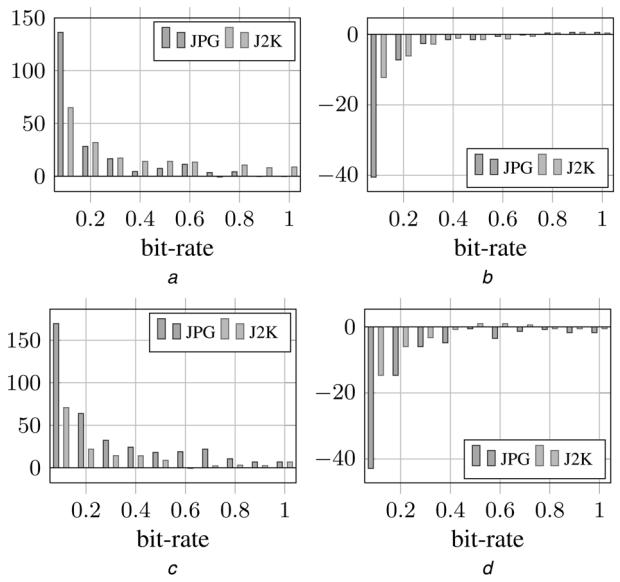

Fig. 6 *Relative change in EERs and IRs compared with the baseline performance for $\mathcal{S}_2$ at different bit rates for LPQ and BSIF*

a ⇅ EER (%) LPQ
b ⇅ IR (%) LPQ
c ⇅ EER (%) BSIF
d ⇅ IR (%) BSIF

Table 4 Detection error for LBP, HOG and HLF for JPG and J2K compression (multiplied by 10^2)

Format	bpp	LBP	HOG	HLF
uncomp.	–	2.49	1.87	1.75
JPG	1.0	2.55	1.88	1.80
	0.9	2.51	1.88	1.80
	0.8	2.48	1.86	1.82
	0.7	2.53	1.82	1.81
	0.6	2.50	1.83	1.92
	0.5	2.53	1.88	1.88
	0.4	2.56	1.92	2.12
	0.3	2.83	1.91	2.71
	0.2	3.32	2.04	3.58
	0.1	4.94	3.11	5.25
J2K	1.0	2.47	1.89	1.86
	0.9	2.51	1.87	1.93
	0.8	2.51	1.92	1.99
	0.7	2.53	2.09	2.21
	0.6	2.68	2.34	2.41
	0.5	2.71	2.64	2.61
	0.4	2.91	2.75	2.80
	0.3	2.84	2.90	2.86
	0.2	2.84	3.18	2.83
	0.1	2.78	3.50	2.93

a positive effect on recognition accuracy [64]. On the other hand, compression algorithms cause distinct artefacts which may result in subject-specific features, that is, image compression may be seen as an effective pre-processing layer prior to the feature extraction stage. For bit-rates down to 0.3 bpp obtained biometric performance is similar for JPG and J2K compression, for lower bit-rates feature descriptors constantly yield better performance of J2K compression (EERs <10%). Due to the fact that identical conditions of biometric input data generally reveal better performance (for considered levels of image quality) it might be beneficial to adjust reference data (available in raw format, cf. $\mathcal{S}_1$) accordingly.

3.4 Ear detection error

In the next scenarios we perform automatic detection of outer ears. Therefore, we first calculate the detection error E from the segmentation result $S(I)$ for an image I with corresponding ground truth mask G_I (both of dimension $m \times n$ pixels), such that for all positions x, y, $G_I[x, y] = 1$ labels 'ear' pixels (otherwise $G_I[x, y] = 0$), as $E = (1/m \cdot n) \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} G_I[x, y] \oplus S(I)[x, y]$. We evaluate the performance of cascaded object detectors with a fixed-size sliding window. We use (i) LBP [65], (ii) HOG [60], and (iii) Haar-like features (HLF) [66]. Detectors are trained with images from WPUTEDB ear database [67] and negative samples from the INIRA person detection dataset [68]. Since the compression rate of a given probe is not known a priori detectors are only trained with good-quality images.

Obtained segmentation errors for considered bit-rates are summarised in Table 4. We observe that ear detection algorithms appear to be rather robust down to bit-rates of 0.3 bpp where HLF and HOG yield lower detection errors than LBP. For lower bit-rates ear detectors work better for J2K-compressed images. Based on obtained results we choose HLF for automated ear detection in subsequent scenarios. In a real-world scenario this choice may further be motivated by the fact that HLF works best on the employed image database in raw format.

3.5 Scenario 3: compressed probe – automatically segmented

The third scenario ($\mathcal{S}_3$) represents an uni-lateral compression where ear detection is performed automatically on the compressed probe image. This scenario is similar to $\mathcal{S}_1$, however, the fully automated system in $\mathcal{S}_3$ is capable of processing large amounts of data, for example, video footage of several surveillance cameras. Thus, $\mathcal{S}_3$ could refer to the case where a large set of subjects has to be compared against a good-quality reference database in order to reveal identities of single subjects, for example, if video footage captured by several cameras installed in public places needs to be analysed.

Table 5 Performance (EER/IR) of LBP, HOG, LPQ, BSIF for JPG, J2K (Scenario 3)

Format	bpp	LBP	HOG	LPQ	BSIF
uncomp.	–	17.61/44.04	23.30/35.01	16.89/49.24	15.26/51.68
JPG	1.0	19.72/38.69	22.52/33.33	16.64/49.08	15.96/48.01
	0.9	19.39/38.07	22.21/33.33	16.65/48.93	16.12/48.17
	0.8	19.23/37.92	22.68/34.71	16.61/50.61	15.18/49.24
	0.7	20.36/33.79	23.24/35.63	16.59/48.93	16.37/47.25
	0.6	22.57/31.50	24.15/32.72	17.91/46.18	16.96/45.57
	0.5	22.87/30.73	24.55/32.72	17.88/48.17	18.48/45.57
	0.4	24.88/25.38	25.05/31.65	18.19/46.94	19.19/41.59
	0.3	30.69/16.36	31.34/22.78	24.74/40.37	25.81/35.93
	0.2	36.83/9.48	37.58/13.15	31.74/29.82	33.91/19.88
	0.1	43.52/2.14	45.36/3.36	42.28/12.39	44.24/5.20
J2K	1.0	20.84/37.46	24.25/34.25	17.62/46.79	16.15/48.01
	0.9	21.58/34.25	24.39/34.40	17.96/47.09	18.15/46.48
	0.8	22.61/31.65	24.40/32.72	18.76/46.94	18.71/44.19
	0.7	25.32/28.44	26.83/30.58	21.62/44.80	21.16/40.98
	0.6	27.29/24.92	30.20/26.15	23.54/40.52	24.14/35.78
	0.5	28.26/22.78	30.28/26.15	22.90/40.52	24.88/35.63
	0.4	31.42/18.96	33.03/23.09	26.75/36.09	28.63/29.51
	0.3	32.39/18.35	33.50/22.17	27.22/35.02	28.21/25.23
	0.2	32.87/14.98	34.33/18.81	26.51/34.25	30.49/22.32
	0.1	34.25/11.47	34.21/15.75	27.36/28.44	31.11/15.75

Performance rates and relative changes in EERs and IRs for LPQ and BSIF obtained for $\mathcal{S}_3$ are shown in Table 5. Compared with $\mathcal{S}_1$ and $\mathcal{S}_2$ where ear segmentation is performed based on a manually annotated ground truth biometric performance significantly drops. We conclude that ear segmentation, which could even fail in case an automated ear detector is applied, represents a crucial and challenging step in the processing chain of an ear recognition system. Further it can be seen that for the vast majority of considered compression rates (down to 0.2 bpp) better performance is obtained for JPG-compressed images. Spurious compression artefacts (i.e. blocking artefacts in high contrast boundary areas) assisting in boundary localisation seem to be the reason for the performance increase of JPG [55]. It is important to note that biometric performance rates obtained in this scenario may still be practical from a forensic point of view. By increasing the amount of returned candidates (e.g. rank-10 instead of rank-1) accuracy is expected to improve significantly.

3.6 Scenario 4: compressed probe and reference – automatically segmented

In the fourth scenario ($\mathcal{S}_4$) reference and probe images are compressed at identical bit-rates representing an uni-lateral

Table 6 Performance (EER/IR) of LBP, HOG, LPQ, BSIF for JPG, J2K (Scenario 4)

Format	bpp	LBP	HOG	LPQ	BSIF
uncomp.	–	13.04/61.92	14.52/61.31	11.08/72.02	10.95/68.04
JPG	1.0	12.55/64.22	13.87/64.37	10.71/71.56	11.57/64.37
	0.9	11.93/63.91	14.38/63.15	10.46/70.95	11.86/64.98
	0.8	12.64/61.77	13.11/64.68	10.18/70.95	11.29/65.75
	0.7	11.92/63.30	13.96/63.30	10.82/69.27	11.26/62.39
	0.6	14.73/58.26	16.62/59.79	12.26/67.43	14.02/61.47
	0.5	13.90/59.17	15.19/60.86	12.02/69.88	12.90/62.99
	0.4	16.16/55.20	16.21/58.72	13.78/66.82	14.92/61.16
	0.3	23.86/45.72	25.62/49.54	22.21/55.20	22.90/49.54
	0.2	33.33/32.87	32.94/33.33	31.22/38.69	31.85/33.94
	0.1	44.03/8.10	43.91/9.02	42.25/9.79	43.76/9.94
J2K	1.0	14.37/59.63	15.98/62.39	12.19/68.96	12.14/64.98
	0.9	14.27/59.79	16.60/60.09	12.80/67.43	13.14/62.39
	0.8	15.05/58.10	15.71/62.54	14.40/66.97	14.37/64.22
	0.7	17.10/56.42	17.41/60.24	15.60/65.14	15.94/61.01
	0.6	20.08/53.98	21.69/55.81	20.18/61.31	20.61/58.72
	0.5	21.88/53.67	22.48/52.45	21.53/57.65	21.09/57.17
	0.4	24.93/46.79	24.96/49.24	24.61/53.82	24.37/52.29
	0.3	24.72/47.55	25.25/49.39	23.55/55.50	24.07/51.99
	0.2	25.61/48.47	26.05/48.32	24.93/53.67	24.16/51.22
	0.1	25.17/43.27	24.25/46.63	21.63/49.24	23.82/48.47

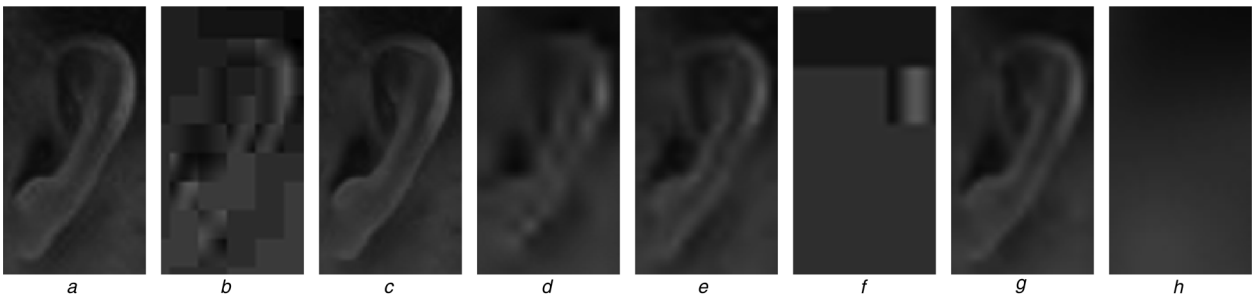

Fig. 7 *Sample close-ups of RoIs of re-scaled and compressed image ID 02463d002 with scaling factors of 50% (a)–(d) and 25% (e)–(h)*

a JPG 1 bpp
b JPG 0.1 bpp
c J2K 1 bpp
d J2K 0.1 bpp
e JPG 1 bpp
f JPG 0.1 bpp
g J2K 1 bpp
h J2K 0.1 bpp

compression scenario. In contrast to $\mathcal{S}_2$, in $\mathcal{S}_4$ ear detection is performed automatically on both images to be compared. This scenario may refer to the case where subjects have to be verified across a large amount of available data captured by similar devices where no good-quality reference is available. From a forensic perspective the automatic tracking of subjects across a large amount of video footage captured by different surveillance cameras could represent a relevant use case of $\mathcal{S}_4$.

Table 6 summarises obtained performance rates for $\mathcal{S}_4$. Interestingly, obtained biometric performance is significantly better across all considered bit-rates compared with $\mathcal{S}_3$. From the baseline performance we observe that while segmentation has been identified as essential processing step the type of segmentation plays an important role as well. In this bi-lateral scenario automated detection is performed on both, the reference and probe image. Further, as mentioned earlier, identical image conditions may additionally improve recognition rates. Again, (such as in $\mathcal{S}_3$) down to bit-rates of 0.3 bpp better performance is obtained for JPG-compressed images which retain clear edges.

4 Further investigations

In addition, we evaluate effects of image scale as well as alternative compression algorithms. We limit these evaluations on the scenario considered as most relevant for forensic applications, that is, $\mathcal{S}_1$.

Table 7 Performance (EER/IR) for different scaling factors using LPQ (Scenario 1)

Format	bpp	100%	50%	25%
uncomp.	–	4.36/82.26	4.31/81.65	5.01/79.97
JPG	1.0	4.75/82.11	4.65/79.66	9.59/62.69
	0.9	4.75/82.11	4.85/79.51	9.34/61.62
	0.8	4.51/82.42	5.03/80.28	10.42/60.55
	0.7	4.70/82.11	4.92/79.82	11.06/56.88
	0.6	4.71/81.80	5.09/80.58	12.74/49.85
	0.5	4.62/82.72	5.38/80.43	13.97/44.50
	0.4	4.44/81.19	5.46/79.66	18.51/32.42
	0.3	4.90/81.96	5.73/77.52	24.58/23.85
	0.2	5.05/81.50	7.44/68.35	33.34/7.03
	0.1	6.87/73.55	19.12/30.28	45.00/2.29
J2K	1.0	4.70/81.65	4.63/80.28	6.31/74.62
	0.9	4.62/81.50	4.75/80.43	6.69/73.39
	0.8	4.59/81.50	4.58/80.89	7.55/72.32
	0.7	4.65/81.96	5.01/79.97	8.48/66.97
	0.6	4.66/82.11	5.16/79.51	9.45/60.40
	0.5	4.63/81.65	5.66/80.12	11.94/51.99
	0.4	4.63/82.11	5.89/78.13	15.26/39.45
	0.3	4.92/80.73	5.93/75.54	22.46/25.69
	0.2	4.90/80.12	7.72/66.06	31.45/12.08
	0.1	5.56/79.05	16.00/33.33	43.19/1.99

4.1 Influence of distance/resolution

On the basis of the theoretical analysis of Section 2.2 sizes of outer ears within original images correspond to an average distance of approximately 2 m between subject and camera [39]. In order to estimate the influence of image compression for larger distances we re-scale original images using scaling factors of 50 and 25%. That is, images are re-scaled and compressed in order to simulate distances of approximately 4 and 8 m. It is important to note that a decrease in image scale directly relates to a decrease in image resolution. This means, scaling factors of 50 and 25% correspond to a decreases of image resolution of original images from 640×480 to 320×240 and 160×120 pixels, respectively. Considered scaling factors result in average sizes of ears of approximately 63×48 pixels and 31×24 pixels, respectively.

Fig. 7 shows an example of RoIs comprising the outer ear extracted from re-scaled images at compression rates of 1.0 and 0.1 bpp. As can be seen, at maximum compression rates of 0.1 bpp the outer ear may be highly degraded or even disappear depending on the acquired image scene. Experimental results for different scaling factors are summarised in Table 7 where the best performing feature extractor is applied, that is, LPQ. Relative changes of EERs and IRs at different scaling factors are depicted in Fig. 8. For a scaling factor of 50% LPQ reveals similar results compared with original images down to bit-rates of 0.5 bpp for both types of image compression. In contrast, a scaling factor of 25% causes a significant performance drop across all considered bit-rates where performance at the maximum compression rate (0.1 bpp) almost refers to random guessing, cf. Figs. 7*f* and *h*. We conclude that the distance of a subject, that is, the size of the outer ear within the image, represents a crucial factor in the presence of image compression. While ear recognition remains feasible on uncompressed re-scaled images compression artefacts of JPG and J2K may severely degrade or completely blur the outer ear region.

4.2 Alternative compression algorithms

As alternative image compression algorithms we consider JPEG-XR (ISO/IEC IS 29199-2) [43] as well as the recently proposed BPG Format [44]. JXR which is based on Microsoft's HD Photo is known to produce higher quality than JPG and faster conversion than J2K. In the default configuration the photo overlay/overlap transformation is only applied to high pass coefficients prior to the photo core transformation. We employ Microsoft's JXR reference software for the conversion task and iteratively adjust quantisation levels to achieve a target bit-rate. The BPG format is based on the H.265/HEVC video format and a valid HEVC bitstream can be reconstructed from the BPG in case a non-modifiable hardware decoder is present. The H.265/HEVC subset that comprises BPG was chosen to support a wide variety of features, for example,

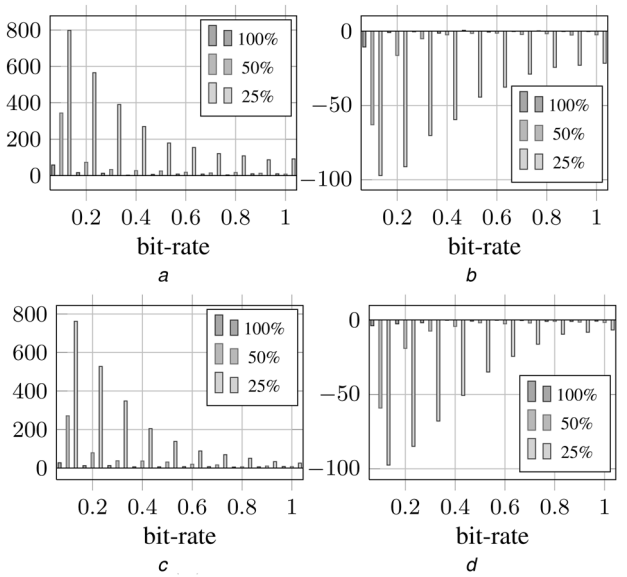

Fig. 8 *Relative change in EERs and IRs compared with the baseline performance for $\mathcal{S}_1$ at different scaling factors for LPQ*

a ⇅ EER (%) JPG
b ⇅ IR (%) JPG
c ⇅ EER (%) J2K
d ⇅ IR (%) J2K

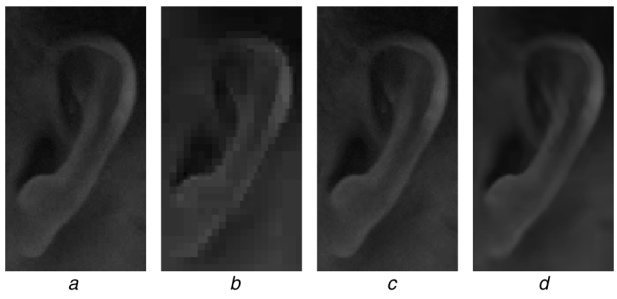

Fig. 9 *Sample close-ups of RoIs of JXR and BPG-compressed image ID 02463d002 at compression rates of 1 and 0.1 bpp*

a JXR 1 bpp
b JXR 0.1 bpp
c BPG 1 bpp
d BPG 0.1 bpp

Table 8 Performance (EER/IR) of LBP, HOG, LPQ, BSIF for JXR, BPG (Scenario 1)

Format	bpp	LBP	HOG	LPQ	BSIF
uncomp.	–	7.03/74.01	8.97/74.77	4.36/82.26	5.05/79.20
JXR	1.0	9.54/62.39	9.39/73.70	4.81/82.11	5.29/78.13
	0.9	10.41/62.39	9.51/70.79	4.88/81.96	5.70/76.91
	0.8	10.99/58.87	9.67/70.18	4.59/81.80	5.77/77.52
	0.7	12.53/52.91	10.38/70.03	4.43/81.04	5.81/76.91
	0.6	14.16/47.25	10.56/67.89	5.05/81.19	5.89/77.22
	0.5	16.14/42.51	10.70/63.76	4.67/81.04	6.07/75.53
	0.4	17.82/35.17	11.88/62.23	4.95/80.73	6.30/73.24
	0.3	20.69/25.84	13.17/54.59	4.95/81.35	7.77/68.65
	0.2	25.56/16.51	15.97/45.57	5.39/80.28	10.50/61.01
	0.1	32.25/6.12	24.57/23.85	6.21/73.70	18.67/35.93
BPG	1.0	8.39/67.74	9.20/72.94	4.44/81.35	6.07/77.06
	0.9	8.90/65.60	9.34/71.87	4.38/81.50	6.12/76.91
	0.8	9.47/64.53	9.41/72.63	4.62/82.26	6.14/76.30
	0.7	9.67/62.534	9.43/72.17	4.67/82.72	6.09/75.08
	0.6	10.31/58.56	9.99/70.64	4.69/82.42	6.41/73.39
	0.5	10.98/53.67	10.58/69.26	4.60/81.19	6.96/71.56
	0.4	11.65/50.46	10.44/69.72	4.90/80.89	7.36/71.71
	0.3	12.12/49.69	10.38/67.43	4.81/81.19	7.90/68.81
	0.2	13.70/45.11	10.90/66.36	5.09/81.35	8.79/64.37
	0.1	16.14/37.77	12.22/57.95	5.46/80.28	10.88/55.05

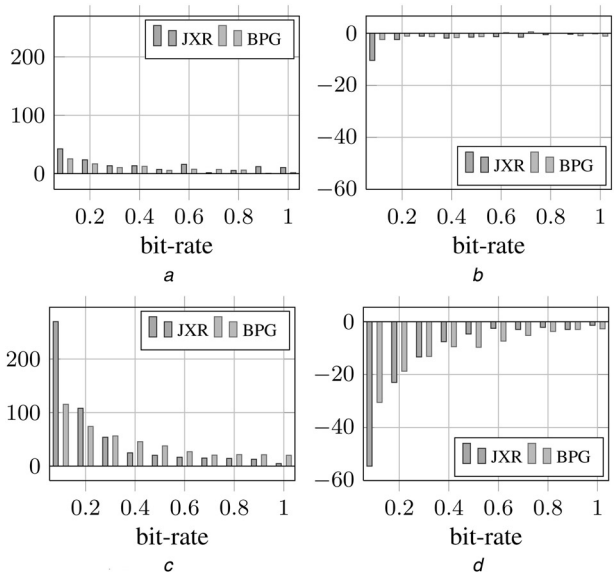

Fig. 10 *Relative change in EERs and IRs compared with the baseline performance for $\mathcal{S}_1$ at different bit rates for LPQ and BSIF*

a ⇅ EER (%) LPQ
b ⇅ IR (%) LPQ
c ⇅ EER (%) BSIF
d ⇅ IR (%) BSIF

animation support, as well support for all features which are present in JPG, for example, colorspaces, as well as the extended JPG standard (JPEG XT) [69], for example, higher dynamic range and lossless compression. The main improvement over the JPEG standard in terms of coding efficiency can be reduced to the smaller block size, combined with an adaptive decomposition quadtree, and intra frame prediction. For BPG compression we iteratively configure quality parameters in order to obtain desired bit-rates using the encoder available at [44].

Examples of resulting RoIs of JXR and BPG-compressed images are shown in Fig. 9. Table 8 shows obtained EERs and IRs for both compression algorithms and Fig. 10 depicts relative changes of EERs and IRs. On the one hand, compared with JPG and J2K best biometric performance is obtained on images compressed with the recently proposed BPG format. The superior rate-distortion performance of BPG compared with J2K or JXR further indicates the feasibility of robust ear recognition at large distances. However, obtained performance gains appear rather small and may not be interpreted as significant. On the other hand, JXR compression does not provide any improvement in terms of biometric performance. We conclude that biometric performance obtained on compressed images rather depends on the employed feature extractor than on the applied image compression technique.

5 Conclusions

In this work we investigated the feasibility and limitations of ear recognition under the influence of image compression which represents an omnipresent type of image distortion in forensic applications. Based on experimental results obtained for three different detection algorithms and four feature extraction methods we conclude that, ear recognition might be feasible in the presence of severe image compression depending on several factors:

(i) The quality of the employed compression algorithm and the compression rate play an important role. For all employed feature extractors as well as detection algorithms we observe differences in biometric performance for the considered compression algorithms. For the majority of compression algorithms bit-rates below 0.3 bpp cause serious drops in biometric performance.

(ii) Ear segmentation is a crucial processing step. Significantly better performance was obtained for scenarios $\mathcal{S}_1$ and $\mathcal{S}_2$ where ears were manually annotated in contrast to the more challenging scenarios $\mathcal{S}_3$ and $\mathcal{S}_4$ employing automated ear detection.
(iii) The type of ear segmentation plays another important role. Better performance was obtained when both, probe and reference image were automatically processed although automatic detectors may perform false segmentations. This finding may call for semi-automated systems where humans approve automatically detected ear regions in case a large amount of data has to be processed.
(iv) The choice of feature extraction is of great importance. At identical compression rates systems relying on distinct feature extractors outperform others in terms of biometric performance across considered compression algorithms.
(v) In contrast to re-scaled uncompressed images, limits of ear recognition on compressed re-scaled images, representing low-resolution scenarios and large distances, are reached relatively fast since compression algorithms tend to cause severe degradations on small regions comprising the outer ear of a subject.

On the basis of the above observations different approaches might be subject to future investigations. While this work considers a NN-classifier and cosine distance at comparison stage, more sophisticated classifier which might depend less on applied feature extraction, for example, sparse representation classification [12, 70, 71], could help to improve the robustness of an ear recognition system in the presence of image compression. However, a more complex classifier might as well increase the response time of the system if it is operated in identification mode. Moreover, super-resolution techniques might be of particular interest in case video footage is available. These methods could be employed in order to reduce compression artefacts and hence improve the robustness and biometric performance of an ear recognition system [41]. Further, feature fusion approaches [9, 11, 72] could be utilised to bridge the observed performance gap between systems operating on raw and highly compressed images.

Finally, it is important to put emphasis on the fact that experimental evaluations have not been conducted on real surveillance footage. However, real surveillance material may not be available in un-compressed form, that is, double compression effects and a non-existent baseline may hinder a thorough investigation of image compression effects on ear recognition. Moreover, the usage of publicly available ear datasets enables fully reproducible research. Therefore, this work represents a fully reproducible study on the influence of image compression techniques on ear biometrics which may serve as an adequate reference point for forensic investigators as well as for operators choosing system settings and as a guidance for experts standardising biometric data interchange formats.

6 Acknowledgment

This work was funded by the Federal Ministry of Education and Research (BMBF) of Germany and the Center for Advanced Security Research Darmstadt (CASED).

7 References

1 Abaza, A., Ross, A., Hebert, C., *et al.*: 'A survey on ear biometrics' (ACM Computer Surveillance, 2013)
2 Chen, H., Bhanu, B.: 'Shape model-based 3d ear detection from side face range images'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition – Workshops, 2005. CVPR Workshops, June 2005, pp. 122–127
3 Abdel-Mottaleb, M., Zhou, J.: 'Human ear recognition from face profile images'. Proc. of the 2006 Int. Conf. on Advances in Biometrics, ser. ICB'06, Springer-Verlag, Berlin, Heidelberg, 2006, pp. 786–792
4 Yan, P., Bowyer, K.: 'Biometric recognition using 3d ear shape', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2007, **29**, (8), pp. 1297–1308
5 Pflug, A., Winterstein, A., Busch, C.: 'Robust localization of ears by feature level fusion and context information'. Proc. of the Int. Conf. on Biometrics (ICB), 2013
6 Wang, Y., Mu, Z.-C., Zeng, H.: 'Block-based and multi-resolution methods for ear recognition using wavelet transform and uniform local binary patterns'. 19th Int. Conf. on Pattern Recognition, 2008. ICPR 2008. December 2008, pp. 1–4
7 Bustard, J., Nixon, M.: 'Robust 2d ear registration and recognition based on sift point matching'. Second IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. September 2008, pp. 1–6
8 Benzaoui, A., Hadid, A., Boukrouche, A.: 'Ear biometric recognition using local texture descriptors', *J. Electron. Imaging*, 2014, **23**, (5), pp. 1–12
9 Chang, K., Bowyer, K., Sarkar, S., *et al.*: 'Comparison and combination of ear and face images in appearance-based biometrics', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2003, **25**, (9), pp. 1160–1165
10 Zhang, H.-J., Mu, Z.-C., Qu, W., *et al.*: 'A novel approach for ear recognition based on ICA and RBF network'. Proc. of 2005 Int. Conf. on Machine Learning and Cybernetics, 2005, August 2005, vol. 7, pp. 4511–4515
11 Nanni, L., Lumini, A.: 'Fusion of color spaces for ear authentication', *Pattern Recogn.*, 2009, **42**, (9), pp. 1906–1913. Available at http://www.sciencedirect.com/science/article/pii/S003132030800455X
12 Khorsandi, R., Cadavid, S., Abdel-Mottaleb, M.: 'Ear recognition via sparse representation and gabor filters'. Fifth Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS'12), 2012, pp. 278–282
13 Pflug, A., Busch, C.: 'Ear biometrics: a survey of detection, feature extraction and recognition methods', *IET Biometrics*, 2012, **1**, pp. 114–129
14 Jain, A.K., Maltoni, D.: 'Handbook of fingerprint recognition' (Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003)
15 Burge, M.J., Bowyer, K.W.: 'Handbook of Iris recognition' (Springer Publishing Company, Incorporated, 2013)
16 Jain, A.K., Li, S.Z.: 'Handbook of face recognition' (Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005)
17 Hoogstrate, A.J., Heuvel, H.V.D., Huyben, E.: 'Ear identification based on Surveillance Camera Images', *Sci. Justice*, 2001, **41**, pp. 167–172
18 News, N.D.: 'Mezuzah arsonist snagged by an ear thanks to facial recognition technology'. 11 April 2013
19 I. O. for Standardization: 'ISO/IEC 19794:2011. Information technology–Biometric data interchange formats', 2011
20 Kasaei, S., Deriche, M., Boashash, B.: 'A novel fingerprint image compression technique using wavelets packets and pyramid lattice vector quantization', *Trans. Img. Proc.*, 2002, **11**, (12), pp. 1365–1378. Available at http://dx.doi.org/10.1109/TIP.2002.802534
21 Funk, W., Arnold, M., Busch, C., *et al.*: 'Evaluation of image compression algorithms for fingerprint and face recognition systems'. IEEE Systems, Man Cybernetics (SMC): Information Assurance Workshop, 2005, pp. 72–78
22 Mascher-Kampfer, A., Stögner, H., Uhl, A.: 'Comparison of compression algorithms impact on fingerprint and face recognition accuracy'. Proc. of SPIE Visual Communications and Image Processing (VCIP'07), 2006, pp. 1–12
23 Elad, M., Goldenberg, R., Kimmel, R.: 'Low bit-rate compression of facial images', *IEEE Trans. Image Process.*, 2007, **16**, (9), pp. 2379–2383
24 Bryt, O., Elad, M.: 'Compression of facial images using the *k*-svd algorithm', *J. Vis. Commun. Image Represent.*, 2008, **19**, (4), pp. 270–282
25 Eickeler, S., Muller, S., Rigoll, G.: 'High quality face recognition in jpeg compressed images'. Proc. 1999 Int. Conf. on Image Processing, ICIP'99, 1999, vol. 1, pp. 672–676
26 Delac, K., Grgic, S., Grgic, M.: 'Recent advances in face recognition', in Delac, K., Grgic, M., Bartlett, M.S. (Eds.): 'ch. Image compression in face recognition – a literature survey' (InTech, 2008) pp. 1–14
27 Daugman, J., Downing, C.: 'Effect of severe image compression on iris recognition performance', *IEEE Trans. Inf. Forensics Sec.*, 2008, **3**, pp. 52–61
28 Ives, R.W., Broussard, R.P., Kennell, L.R., *et al.*: 'Effects of image compression on iris recognition system performance', *J. Electron. Imaging*, 2008, **17**, pp. 011015-1-011015-7
29 Grother, P.: 'Quantitative standardization of iris image formats'. Proc. of the Biometrics and Electronic Signatures (BIOSIG 2009), ser. LNI, 2009, pp. 143–154
30 Ives, R.W., Bishop, D.A., Du, Y., *et al.*: 'Iris recognition: The consequences of image compression', *EURASIP J. Adv. Signal Process.*, 2010, **2010**, pp. 1–9
31 I. O. for Standardization: 'ISoO/IEC IS 15444-1:2004.Information technology–JPEG 2000 image coding system: Core coding system', 2004
32 Orandi, S., Libert, J., Grantham, J., *et al.*: 'NIST Special Publication 500 – 289: Compression Guidance for 1000 ppi Friction Ridge Imagery'. February 2014
33 Rathgeb, C., Uhl, A., Wild, P.: 'Iris Recognition: From Segmentation to Template Security', ser. Advances in Information Security (Springer, New York, 2012), vol. 59, ch. 8 (Image Compression Impact on Iris Recognition), pp. 97–140
34 Orandi, S., Libert, J.M., Grantham, J.D., *et al.*: 'NISTIR 7780: Effects of JPEG 2000 Lossy Image Compression on 1000 ppi Latent Fingerprint Casework'. 2013
35 Jeong, G.-M., Kim, C., Ahn, H.-S., *et al.*: 'JPEG quantization table design for face images and its application to face recognition', *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 2006, **E-69**, (11), pp. 2990–2993
36 Chen, M., Zhang, S., Karim, M.: 'Modification of standard image compression methods for correlation-based pattern recognition', *Opt. Eng.*, 2004, **43**, (8), pp. 1723–1730
37 Konrad, M., Stögner, H., Uhl, A.: 'Evolutionary optimization of JPEG quant tables for compressing iris polar images in iris recognition systems'. Proc. of the sixth Int. Symp. on Image and Signal Processing and Analysis, ISPA '09, Salzburg, Austria, September 2009
38 Hämmerle-Uhl, J., Karnutsch, M., Uhl, A.: 'Evolutionary optimisation of JPEG2000 Part 2 wavelet packet structures for polar iris image compression'. Proc. 18th Iberoamerican Con. on Pattern Recognition (CIARP'13), ser. LNCS, 2013, vol. 8258, pp. 391–398
39 Wagner, J., Pflug, A., Rathgeb, C., *et al.*: 'Effects of severe signal degradation on ear detection'. Proc. of IEEE Int. Workshop on Biometrics and Forensics (IWBF). March 2014

40 Pflug, A., Wagner, J., Rathgeb, C., *et al*.: 'Impact of severe signal degradation on ear recognition performance'. Proc. of Biometrics and Forensics and De-identification and Privacy Protection (BiForD), 2014
41 Watabe, D., Minamidani, T., Zhao, W., *et al*.: 'Effect of barrel distortion and super-resolution for single-view-based ear biometrics rotated in depth'. Int. Conf. on Biometrics and Kansei Engineering (ICBAKE'13), 2013, pp. 183–188
42 I. O. for Standardization: 'ISO/IEC IS 10918-1:1994. Information technology–Digital compression and coding of continuous-tone still images: requirements and guidelines', 1994
43 I. O. for Standardization: 'ISO/IEC IS 29199-2:2012. Information technology–JPEG XR image coding system – Part 2: Image coding specification', 2012
44 Bellard, F.: 'BPG image format'. Available at http://bellard.org/bpg, retrieved January 2015
45 'ImageMagick Software'. Available at http://www.imagemagick.org/download/, retrieved April 2015
46 Engel, D.: 'JJ 2000', available at http://www.wavelab.at/sources/, retrieved April 2015
47 Yan, P., Bowyer, K.W.: 'Biometric recognition using 3D ear shape', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2007, **29**, pp. 1297–1308
48 Budagavi, M., Fuldseth, A., Bjontegaard, G., *et al*.: 'Core Transform Design for the High Efficiency Video Coding (HEVC) Standard', *IEEE J. Sel. Top. Signal Process.*, 2013, **7**, (6), pp. 1029–1041
49 Faltemier, T.C., Bowyer, K.W., Flynn, P.J.: 'Rotated Profile Signatures for robust 3D Feature Detection'. Automatic Face and Gesture Recognition, 2008
50 Watabe, D., Sai, H., Sakai, K., *et al*.: 'Improving the robustness of single-view ear-based recognition under a rotated in depth perspective'. Int. Conf. on Biometrics and Kansei Engineering (ICBAKE'11), 2011, pp. 179–184
51 Sforza, C., Grandi, G., Binelli, M., *et al*.: 'Age- and sex-related changes in the normal human ear', *Forensic Sci. Int.*, 2009, **2009**, pp. 110e1–110e7
52 ISO/IEC TC JTC1 SC37 Biometrics: 'ISO/IEC 19795-1:2006 Information techn–biometric performance testing and reporting–Part 1: principles and framework' (Int'l Organization for Standardization and Int'l Electrotechnical Committee, March 2006)
53 Wang, Z., Bovik, A., Sheikh, H., *et al*.: 'Image quality assessment: from error visibility to structural similarity', *IEEE Trans. Image Proc.*, 2004, **13**, (4), pp. 600–612
54 Thomos, N., Boulgouris, N., Strintzis, M.: 'Optimized transmission of jpeg2000 streams over wireless channels', *Trans. Image Process.*, 2006, **15**, (1), pp. 54–67
55 Rathgeb, C., Uhl, A., Wild, P.: 'Effects of severe image compression on iris segmentation performance'. Proc. of the Int. Joint Conf. on Biometrics (IJCB14), 2014, pp. 1–6
56 Ahonen, T., Hadid, A., Pietikainen, M.: 'Face description with local binary patterns: Application to face recognition', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2006, **28**, (12), pp. 2037–2041
57 Abaza, A., Harrison, M.A.F.: 'Ear recognition: a complete system', *SPIE 8712, Biometric Surveillance Technol. Hum. Activity Identif.*, 2013, **8712**, pp. 1–11
58 Ahonen, T., Rahtu, E., Ojansivu, V., *et al*.: 'Recognition of blurred faces using local phase quantization'. Int. Conf. on Pattern Recognition, December 2008, pp. 1–4
59 Ojansivu, V., Heikkila, J.: 'Blur insensitive texture classification using local phase quantization' in 'Image and Signal Processing' (Springer, Berlin Heidelberg, 2008), (ser. *Lecture Notes in Computer Science*), pp. 236–243
60 Dalal, N., Triggs, B.: 'Histograms of oriented gradients for human detection'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, CVPR, June 2005, vol. 1, pp. 886–893
61 Kannala, J., Rahtu, E.: 'Bsif: Binarized statistical image features'. IEEE Int. Conf. on Pattern Recognition, 2012, pp. 1363–1366
62 Raghavendra, R., Busch, C.: 'Texture based features for robust palmprint recognition: a comparative study', *EURASIP J. Inf. Sec.*, 2015, **2015**, pp. 1–9
63 Štruc, V., Pavešić, N.: 'The complete gabor-fisher classifier for robust face recognition', *EURASIP J. Adv. Signal Process*, 2010, **2010**, pp. 31:1–31:13
64 Rakshit, S., Monro, D.M.: 'An evaluation of image sampling and compression for human iris recognition', *IEEE Trans. Inf. Forensics Sec.*, 2007, **2**, pp. 605–612
65 Ojala, T., Pietikainen, M., Maenpaa, T.: 'Multiresolution gray-scale and rotation invariant texture classification with local binary patterns', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2002, **24**, (7), pp. 971–987
66 Viola, P., Jones, M.: 'Rapid object detection using a boosted cascade of simple features'. Proc. of the 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, CVPR 2001, vol. 1, pp. 511–518
67 Frejlichowski, D., Tyszkiewicz, N.: 'The West Pomeranian University of Technology Ear Database A Tool for Testing Biometric Algorithms' Image Analysis and Recognition, (Springer, 2010)
68 Dalal, N., Triggs, B.: 'Histograms of oriented gradients for human detection'. Int. Conf. on Computer Vision and Pattern Recognition, 2005, vol. 2, pp. 886–893
69 Richter, T.: 'On the standardization of the JPEG XT image compression'. Picture Coding Symp. (PCS), 2013, December 2013, pp. 37–40
70 Wright, J., Ma, Y., Mairal, J., *et al*.: 'Sparse representation for computer vision and pattern recognition', *Proc. IEEE*, 2010, **98**, (6), pp. 1031–1044
71 Kumar, A., Chan, T.-S.T.: 'Robust ear identification using sparse representation of local texture descriptors', *Pattern Recogn.*, 2013, **46**, (1), pp. 73–85
72 Yuan, L., chun Mu, Z.: 'Ear recognition based on local information fusion', *Pattern Recogn. Lett.*, 2012, **33**, (2), pp. 182–190