Recompression effects in iris segmentation

Anonymous ICB 2015 submission

Abstract

The ABSTRACT is to be in fully-justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word "Abstract" as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. Leave two blank lines after the Abstract, then begin the main text. Look at previous ICB abstracts to get a feel for style and length.

1. Introduction

Iris recognition [3, 24] is one of the most deployed biometric modalities, standardized by the International Civil Aviation Organization (ICAO) for use in future passports, and one of the technologies in the Unique Identification Authority of India's (UID) Aadhaar project to uniquely identify Indian citicens. The increasing market saturation of biometric instead of conventional access control methods raises the need for efficient means to store such data. The International Organization for Standardization (ISO) specifies iris biometric data to be recorded and stored in (raw) image form (ISO/IEC FDIS 19794-6) rather than in extracted templates (e.g. iris-codes). Such deployments benefit from future improvements (e.g. in feature extraction stage) which can be easily incorporated without re-enrollment of registered users. Since biometric templates may depend on patent-registered algorithms, databases of raw images also enable more interoperability and vendor neutrality [24]. These facts motivate detailed investigations and optimisations of image compression on iris biometrics in order to provide an efficient storage and rapid transmission of raw biometric records. Furthermore, the application of low-powered mobile sensors for image acquisition, e.g. mobile phones, raises the need for reducing the amount of transmitted data.

The certainly most relevant standard for compressing image data relevant in biometric systems is the ISO/IEC 19794 standard on Biometric Data Interchange Formats, where in the most recently published version (ISO/IEC FDIS 19794-6), only JPEG2000 is included for lossy compression. JPEG2000 has also been recommended for various application scenarios and standardised iris images (IREX records) by the NIST Iris Exchange[1] program. The ANSI/NIST-ITL 1-2011 standard on "Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information" (former ANSI/NIST-ITL 1-2007) also supports only JPEG2000 for applications tolerating lossy compression.

As a consequence, according to the importance of this issue, many studies comparing and optimising lossy compression techniques for iris imagery may be found in the literature. Since the CASIA iris datasets have been very popular among researchers ever since their establishment, many papers dealing with compression have been relying on the (extended) CASIA V1.0 dataset, including also first IREX investigations [22, 14, 17, 7, 15] (apart from other examples using the ICE 2005 dataset [6, 13]).

Since it has been pointed out [21] that the CASIA V1.0 dataset exhibits manipulated pupil areas and should therefore not be used any further in experimentation, compression researchers moved to other (and more recent, more challenging etc.) datasets, e.g. the CASIA V3.0 [9, 24], the CASIA V4.0 [27], the Bath [14, 19], and the UBIRIS.v1 [7, **?**] datasets. While the images of CASIA V1.0 and ICE 2005 are given in uncompressed format, images in CASIA V3.0, CASIA V4.0, UBIRIS and Bath datasets are provided as JPEG (the first three) or JPEG2000 (the latter) lossy compressed data. Therefore, any compression experiments conducted on these datasets operate on pre-compressed data.

This fact has not been ignored entirely – for example, in [24], preparatory JPEG compression experiments with uncompressed data reveal that slightly pre-compressed data leads to better recognition performance due to denoising effects. Thus experiments with pre-compressed data are assessed to be unproblematic. The same argument is used for JPEG2000 pre-compressed data [19] based on the results in [14]. However, eventual artifacts resulting from recompression effects are not accounted for in these considerations. Recompression artifacts arise in cases where data is compressed twice (or multiple times) with lossy compression schemes, i.e. where artifacts from the first compression step (termed pre-compression) are aggravated or exploited

[1] IREX `http://iris.nist.gov/irex/`

by the second compression step.

Two different types of such effects may be distinguished: First, *intra-recompression*, where the same compression scheme is used several times, whereas in *inter-recompression* different methods are used in the different compression steps. For example, using JPEG pre-compressed data and applying JPEG XR and JPEG2000 [9] or JPEG2000 and fractal compression [**?**] is eventually prone to inter-recompression artifacts, while the application of JPEG to JPEG pre-compressed data [27, 24] can be prone to intra-recompression artifacts.

While next to nothing can be found on the issue of inter-recompression artifacts in the general compression literature, intra-recompression artifacts are better investigated, at least in the case of lossy JPEG compression. Soon after the establishment of the JPEG standard [20], it was found that JPEG recompression artifacts arise and do not follow a linear behaviour [4]. Extensive experiments in this direction can also be found in [**?**], and following these obversations, requantisation-based schemes have been suggested for JPEG, reducing recompression artifacts considerably [1]. Recently, the identification of images which underwent JPEG double compression (i.e. JPEG intra-recompression) has been a hot topic in image forensics [26].

Taking all these facts together, it gets clear that recompression artifacts may impact experimental results with respect to biometric recognition performance, an issue, that has been neglected so far. As discussed, ISO/IEC FDIS 19794-6 requires storing biometric data as raw images, hence all components of a biometric system are affected when operating with compressed data. As we investigate the recompression issue by studying impact on an iris recognition system, the influence on segmentation and texture extraction as well as feature extraction, i.e. iris code computation, has to be evaluated. [25, 6] suggest that data reduction has the highest impact on the iris segmentation. Since segmentation is also the first step in the pipeline, this potentially effects the performance of later steps as well and is therefore of particular importance.

We systematically investigate eventual intra- and inter-recompression effects in an experimental study for iris segmentation. Given the importance of JPEG in the area (as the CASIA V3.0/V4.0 and UBIRIS.v1 datasets are only available in this format), we focus on JPEG pre-compressed data. In our experiments, we compare iris segmentation and general purpose image quality metrics applied to single compressed vs. recompressed (i.e. JPEG pre-compressed) iris image data. In Section 2, we discuss the aspects relevant when generating single- and recompressed data. Under these aspects, data for experimental evaluation is generated as described in section 3. Section 4 describes and lists results of several experiments, which are compared in section **??**. From the experiments' individual and comparison results, we withdraw a conclusion in section 6.

2. Compression scheme

As discussed, we investigate whether there is a difference in using truly uncompressed data or pre-compressed data in experiments rating the performance of an iris segmentation. Using pre-compressed data in such experiments means a pre-compressed image I_p is compressed a second time, resulting in an recompressed image I_r. When compressing a truly uncompressed image I_u, the resulting image I_s is generated in a single compression step. Since experiments are typically carried out on a data set with more than one image, we denote $I_u^{(k)}, I_p^{(k)}, I_s^{(k)}, I_r^{(k)} \in \mathbb{R}^{w \times h}$ as the k^{th} image with width w and height h. For simplicity, I_u, I_p, I_r, I_s subsequently denote a particular image but unspecified image of a data set. Furthermore, we define $s(F) \in \mathbb{N}$ with $F \sim I \in \mathbb{R}^{w \times h}$ as a function that returns the file size of the file F storing an image I. Since common lossy compression algorithms also employ lossless compression methods, e.g. run-length encoding, before writing to a file, F is only loosely related to the pixel data, namely the image, I. For simplicity, we denote $s(I)$ as the file size of the file F encoding the pixel values of an image I. $c_m(I, q)$ with $q \in \mathbb{N}$ describes the process of compressing an image I using a particular method m parametrized with the quality parameter p. In terms of this paper we use the values $m \in \{jpg, jxr, j2k\}$, where

- jpg corresponds to the well-known (ISO/IEC IS 10918-1) DCT-based image compression method JPEG [10],
- $j2k$ corresponds to the wavelet-based image compression standard JPEG2000 (ISO/IEC IS 15444-1), which can operate at higher compression ratios [11] and
- jxr corresponds to a compression standard JPEG-XR based on Microsofts HD Photo, which is specified in (ISO/IEC IS 29199-2) [12].

For rating an image I's compression effectiveness, we define the compression ratio cr between an uncompressed image I_u and a compressed image I_c as

$$cr(I_u, I_c) = \frac{s(I_u)}{s(I_c)} \quad \text{with} \quad I_c \in \{I_r, I_s\} \tag{1}$$

For the later described experiments, images are compressed to a target compression ratio $cr_t \in \mathbb{R}$. However, only the *j2k* compression standard [11] allows to specify a target compression ratio cr_t directly via parameter q. Hence this is the only method where we can control the file size $s(I_c)$ directly. The other two compression methods take a quality parameter $q \in \mathbb{N}$ only, which controls the quality but not the file size of the output image I_c. Therefore it is not

possible to set this parameter in a way to meet a certain target compression ratio cr_t. Due to the limited universe of the quality parameters, the target compression ratio cr_t cannot be achieved exactly for any of the three methods. Parameter optimisation can be done, such that $cr_t \approxeq cr(I_u^{(k)}, I_c^{(k)})$. We propose an algorithm to compress a set of K uncompressed images I_u using a particular method m to achieve a certain compression ratio cr_t in a way that the compression ratio of each image is met as close as possible. This process, illustrated in Fig. 1, employs

1. Compute the single-compressed image $I_s^{(k)}$ with method m such that $cr(I_u^{(k)}, I_s^{(k)}) \approx cr_t$. The optimal quality parameter $q_s^{(k)}$ is computed for each image separately by

$$s_t^{(k)} = \frac{s(I_u^{(k)})}{cr_t} \tag{2}$$

$$q_s^{(k)} = \underset{q\in\mathbb{N}}{argmin} |s(c_m(I_u^{(k)}, q)) - s_t^{(k)}|, \tag{3}$$

where $s_t^{(k)}$ is the file size exactly meeting the target compression ratio cr_t. This is implemented by iteratively searching the quality parameter q that results in the closest achievable compression ratio $cr(I_u, I_c)$. The single compressed images $I_s^{(k)}$ using method m are computed with the optimal parameters $q_s^{(k)}$ as

$$I_s^{(k)} = c_m(I_u^{(k)}, q_s^{(k)}) \tag{4}$$

2. Compute a pre-compressed image $I_p^{(k)}$ representing images in pre-compressed data set using jpg-method [10] with an arbitrary but fixed quality parameter q_p, i.e.

$$I_p^{(k)} = c_{jpg}(I_u^{(k)}, q_p) \tag{5}$$

3. Now, find a quality parameter $q_d^{(k)}$ that allows to compress the pre-compressed image $I_p^{(k)}$ a second time, such that the resulting recompressed image $I_r^{(k)}$ has the same file size as the single-compressed image $I_s^{(k)}$, i.e. $s(I_s^{(k)}) \approxeq s(I_r^{(k)})$. Such a quality parameter $q_r^{(k)}$ can be found by optimising

$$q_r^{(k)} = \underset{q\in\mathbb{N}}{argmin} |s(c_m(I_p^{(k)}, q)) - s(I_s^{(k)})| \quad \forall s(I_s^{(k)}) \geq s(I_r^{(k)}) \tag{6}$$

The condition $s(I_s^{(k)}) \geq s(I_r^{(k)})$ is of importance to establish fair conditions, since it is very likely that the file sizes $s(I_s^{(k)}), s(I_r^{(k)})$ cannot be equalized due to the limited universe of the quality parameters q. The recompressed images $I_r^{(k)}$ are then computed from the pre-compressed images $I_p^{(k)}$ with the found optimal parameters q_r as

$$I_r^{(k)} = c_m(I_p^{(k)}, q_r^{(k)}) \tag{7}$$

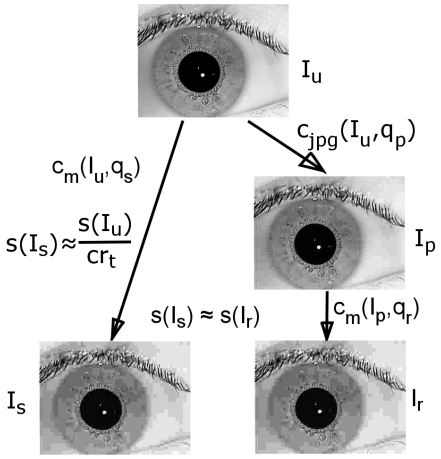


Figure 1. Basic compression principle to obtain two images achieving approximately the same target compression ratio cr_t from an uncompressed image $I_u^{(k)}$ using a particular compression method m. One image, $I_s^{(k)}$, is compressed in a single step while the other, $I_r^{(k)}$, uses a pre-compression and a final compression step. Note the pre-compression step is always a jpg-compression, while the final compression step uses the same method m as used for single-compression.

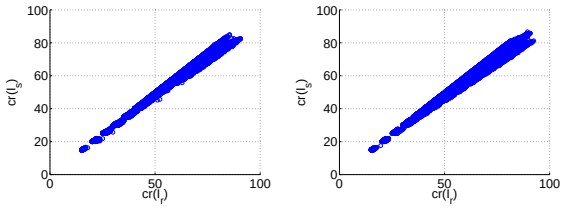
Using data sets generated with this method, we can investigate the impact of artifacts in an recompressed image I_r in comparison to those in single-images I_s. The first one contains artifacts by two compression, while the latter one contains artifacts from one compression step only.

3. Experimental setup

Although there are several iris data sets around, few are available in uncompressed format. We use the IITD Iris data base[2]. The main reason for this is the availability of a segmentation ground truth created by an expert, which was recently introduced by Hofbauer *et al.* [8] and used in [25]. The k^{th} image of this segmentation ground truth data set is subsequently denoted as $SGT^{(k)}$. According to information by the IITD iris data base's authors, the images, stored in a 3-channel uncompressed bitmap format[3] are already JPEG-compressed with 100% quality by the sensor (JIRIS, JPC1000). Since they are stored as bitmaps, all images have an identical file size of $s(I_u)$=230,454 bytes. Despite not being optimal, using the IITD was necessary due to the available ground truth, for reasons becoming obvious in section 4.2. Futhermore, the IITD – contrary to others, e.g. the ND-IRIS-0405 iris image dataset [2] – is captured

[2]IITD Iris Database version 1.0, www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm

[3]We want to point out that storing in 1-channel bitmaps would be more efficient, since the images were captured in near-infrared. However, we use the size information of the 3-channel bitmap in computing compression ratios



	single: $1-\frac{cr(I_u,I_s)}{cr_t}$			recomp.: $1-\frac{cr(I_u,I_r)}{cr_t}$		
%	*jpg*	*j2k*	*jxr*	*jpg*	*j2k*	*jxr*
μ	-3.21	-3.48	-4.33	-6.80	-7.04	-9.49
σ	2.74	2.51	2.87	4.83	4.23	4.34

Figure 2. Scatter plots of measured compression ratios $cr(I_u, I_s)$ over $cr(I_u, I_r)$ for methods *jpg* (left) and *jxr* (right). The graphs indicate that the $s(I_s^{(k)}) \geq s(I_r^{(k)})$ condition from equ. (6) is satisfied. While this is indeed true for *j2k* and *jxr*, we observe a violation in 0.13% of the cases for *jpg* at $cr_t \geq 70$, because JPEG is already working at it's boundaries at such high compression ratios. The table below reveals, that in average the aimed cr_t is met with 3.67% accuracy for single-compressed images I_s, while the re-compressed ones I_r only reach 7.8%. This is due to the limited universe of quality parameters q, q_p.

under favorable conditions, which allows for lower segmentation errors. This is necessary to distinguish between noise and recompression-effects.

We use the scheme introduced in section 2 to compress obtain data sets with target compression ratios

$$cr_t \in \{15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75\}. \quad (8)$$

For each of these target compression ratios cr_t, the pre-compression step in recompression mode is carried out with quality parameters

$$q_p \in \{100, 80, 75, 70\} \quad (9)$$

to simulate different levels of pre-compression. Each of these combinations is used to compress with the introduced jpg, $j2k$ and jxr methods. We start at compression ratio $cr_t = 15$, because even a pre-compression with $q_p = 100$ achieves - depending on the image's content - already a compression ratio of $cr(I_u, I_p) \approx 10$. For obvious reasons, no smaller compression ratio $cr(I_u, I_r) < cr(I_u, I_p)$ can be reached in recompression. This results in a total of 195 data sets with 2240 images each, whose distribution is shown and discussed in fig. 2.

4. Evaluation

Besides assessing the image quality with fully-referenced metrics (section 4.1), we investigate the behaviour of iris segmentation using a segmentation error rate (section 4.2).The individual results are then compared in **??**.

4.1. Full-referenced quality metrics

Evaluating the quality of the compressed images in respect to the original an assortment of full-reference metrics was chosen. The choice was made according to different aspects of human perception starting from mathematically defined to low-level features based and finally to high-level features based. The following were included:

- PSNR: Peak signal-to-noise ratio
- MSSIM[28]:Multi-scale structural similarity index is an extension of the SSIM metric. After the extraction of luminance, structure and contrast components from the image at scale 1, the algorithm iteratively applies a low pass filter and a downsamples the filtered image by a factor of 2. The overall result is the combination of measurements at different scales.
- NQM[18]:Noise Quality Measure, a low-level HVS features based metric. The contrast pyramid of Pelis work was used to model the variation in contrast, sensitivity with distance, dimensions and spatial frequency of the stimuli, and with the variation of their local luminance mean.
- RFSIM[16]:Riesz-transform feature based similarity metric approximates HVS by perceiving an image mainly according to its low-level features. Uses the 1st-order and 2nd-order Riesz transform coefficients. The similarity index is measured by comparing the two feature maps at key locations marked by the feature mask. The mask is generated by a Canny operator.
- VSNR[5]:Visual Signal-to-Noise Ratio, a wavelet based metric. The metric is designed to evaluate both low-level and mid-level HVS features. VSNR works in two stages: the first computes the contrast detection thresholds, while the second estimates visual fidelity by measuring the perceived contrast and the extent to which the distortions disrupt global precedence

Applying these quality metrics *jpg*, *j2k* and *jxr* resulted in the following observation in fig. 3 & 4 :

1. For *jpg* :
   (a) For $cr_t > 15$ single compressed images were of higher quality compared to re-compressed
   (b) At cr_t=15 the single compressed images were of the lowest quality as shown in fig. 3 for MSSIM
   (c) The quality of the re-compressed images followed the trend that the higher the q_p the better the quality of the image
   (d) Previous observation was unanimous for all metrics

2. For *jk2*:

   (a) For all compression ratios and all the metrics The quality of the re-compressed images followed the trend that the higher the q_p the better the quality of the image

   (b) Single compressed images were of the highest quality compared to re-compressed data, for all metrics and compression ratios as shown for MSSIM in fig. 3

3. For *jxr*:

   (a) For $15 \leq cr_t \leq 40$ single compressed images were of the lowest quality compared to re-compressed data for MSSIM and VSNR. Then the latter followed the trend of the higher the q_p the better the quality of the image fig. 4

   (b) For $45 \leq cr_t \leq 75$ images of single compression became of the highest quality and re-compressed data continued the same trend for MSSIM and VSNR.

   (c) NQM showed the same behaviour, but at cr_t=50 as a changing point

   (d) RFSIM showed a different trend from the previous, here single compressed data was always of best quality and re-compressed data followed in quality.

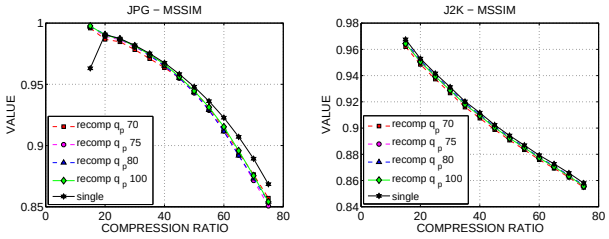


Figure 3. Left: MSSIM of *jpg* single- and recompressed data. Right: MSSIM of *jp2* single- and recompressed data

4.2. Segmentation error rates

In iris recognition, the segmentation of an iris image is considered as one of the most critical parts [6, 25]. We investigate the differences of single- and recompression as well as the aspects of which reference to use. We distinguish between using an absolute reference, e.g. a ground truth, and a relative one, e.g. the segmentation of the uncompressed images I_u, when computing the error rate.

The segmentation accuracy is rated by the mean segmentation error rate, which corresponds to the suggested E1 error rate in the Noisy Iris Challenge Evaluation - Part I

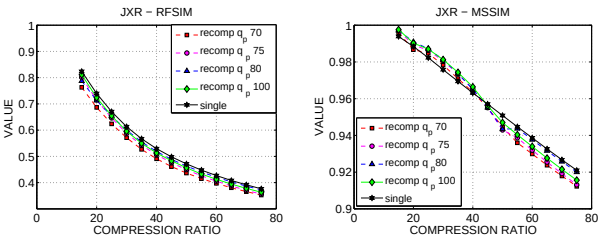


Figure 4. Left: RFSIM of *jxr* single- and recompressed data. Right: MSSIM of *jxr* single- and recompressed data

Figure 5. Segmentation masks of the expert ground truth [25], relative groundtruth $seg(I_u^{(k)})$ and an actual segmentation result $seg(I_r^{(k)})$ (f.l.t.r)

(NICE.I). We define the segmentation error rate ser as

$$ser(R,S) = \overline{R \oplus S} \in [0,1] \quad with \quad R,S \in \{0,1\}^{w \times h}, \tag{10}$$

where R is the binarized reference segmentation and S the binarized segmentation result of the same image I. The mean value of the pixel-wise exclusive-or is the percentage of pixels different in the segmented image S in respect to the reference R. Due to multiple images in a data base, the mean segmentation error $mser$ is computed from K images. We compute the absolute mean segmentation error $mser_{abs}$ in respect to the ground truth SGT and the relative mean segmentation error $mser_{rel}$ in respect to the segmentation of the uncompressed images I_u for single- and recompressed images $I_c \in I_s, I_r$. By denoting the segmentation result of an image I as $seg(I) \in \{0,1\}^{w \times h}$ we have

$$mser_{abs} = \frac{1}{K} \sum_{k=1}^{K} ser(SGT^{(k)}, seg(I_c^{(k)})) \tag{11}$$

$$mser_{rel} = \frac{1}{K} \sum_{k=1}^{K} ser(seg(I_u^{(k)}), seg(I_c^{(k)})) \tag{12}$$

The absolute segmentation error rate is considered to be optimal because of the available ground truth. However, for most data bases no such ground truth is available. Therefore we evaluate if the same conclusions as from the $mser_{abs}$ can be withdrawn from the $mser_{rel}$. The benefit of such a relation (if it exists) is that the $mser_{rel}$ can be computed for any arbitrary data set.

The data set described in section 3 is used to test the two iris segmentation algorithms, Contrast-adjusted Hough Transform (CAHT) and Weighted Adaptive Hough and El-

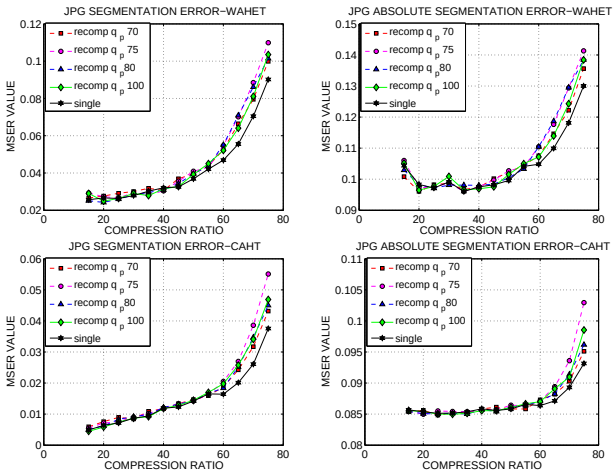


Figure 6. Relative and absolute segmentation error $mser_{rel}$ (left) and $mser_{abs}$ (right) with WAHET (top) and CAHT (bottom) segmentation on *jpg*-compressed data. Note that the $mser_{abs}$ is generally higher than $mser_{rel}$, because the tested algorithms ignore eyelids, yet they are considered in the expert ground truth [25].

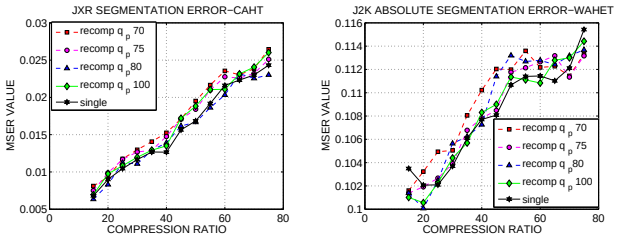


Figure 7. Relative CAHT segmentation error rate $mser_{rel}$ for *jxr*-compressed data (left) and absolute WAHET segmentation error rate $mser_{abs}$ for *jp2*-compressed data (right)

lipsopolar Transform (WAHET), from the USIT Framework v1.0.3[4]. From the results we observe the following:

1. Results for intra-recompression experiments, namely *jpg* on *jpg* pre-compressed data, in figure 6 indicate:

   (a) For small and medium compression ratios ($cr_t \leq 50$) no significant difference in segmentation errors of single- and recompressed data is observeable. This implies that for these compression ratios it has no impact whether pre-compressed or uncompressed data is used in experiments.

   (b) For large compression ratios ($cr_t > 50$), segmentation errors tend to be lower for single-compressed data compared to recompressed data. Thus using pre-compressed or uncompressed data in experiments matters. To be more precise, using recompressed data leads to poorer results as if uncompressed data would have been used.

   (c) $mser_{rel}$ and $mser_{abs}$ generally show the similar trends for medium and large compression ratios, i.e. there is a strong correlation of $mser_{rel}$ and $mser_{abs}$ for $cr_t > 30$. This means, the relative error $mser_{rel}$ suffices to rate performance on iris segmentation here, implying no expert-generated ground truth is needed.

   (d) However, WAHET segmentation errors reveal that in some cases there can be a difference between $mser_{rel}$ and $mser_{abs}$ for $cr_t \leq 30$. Figure 6 shows here a different behaviour between absolute error $mser_{abs}$ (top-right) and relative error $mser_{rel}$ (top-left). Hence, for low compression ratios a ground truth is required to reliably rate a compression method's performace in iris segmentation.

   (e) In recompression, one might expect a linear relation between used pre-compression quality q_p and ranking of the error rates. Interestingly, when looking at figure 6 at high compression ratios, the poorest performance corresponds to $q_p = 75$, while the best is related to $q_p = 70$. In contradiction, $q_p = 100$ performs significantly better than $q_p = 80$ in most settings.

2. There are no clear trends for inter-recompression experiments, namely *jxr* or *j2k* on *jpg* pre-compressed data. Even so, some interesting observations are made, which are illustrated figure 7:

   (a) Data generated in a single compression step generally tends to result in smaller error rates compared to those computed from recompressed data. Interestingly, for extreme values, namely very small and very large compression ratios, single-compression performs often poorer.

   (b) For all experiments carried out with *jxr* and *jp2*, the error rate stagnates in some way for medium compression ratios, i.e. $45 \leq cr_t \leq 70$. Exemplarily this can be seen in the $mser_{abs}$ for *jp2*-method (figure 7 left). The characteristics of a curve's stagnancy vary depending on the pre-compression quality q_p. Since stagnancy can be seen in single- as well as recompressed data, we conclude the effect is generally related to the used methods *jxr* and *jp2*. However, the characteristics of the stagnancy seem be controlled by the pre-compression quality q_p in a way that the lower the pre-compression quality is, the clearer the curve stagnates.

In intra-recompression, recompression effects have a strong impact on experimental results for large compression

[4] as available at `http://wavelab.at/sources/` [23]

ratios, i.e. $cr_t > 50$ (1a, 1b). Researchers are often forced to use precompressed data sets for the sake of ground truth availability. Results for compression ratios of $cr_t > 50$ can therefore not be considered entirely reliable (1b). However, recompression effects have negative influence on segmentation error rates, hence by using uncompressed data for the same experiments, better results may be achieved. If this behaviour is related to intra-recompression in general or for *jpg*-recompression only, is topic to further research.

From 1c we know that for large compression ratios, i.e. $cr_t > 50$, there is no difference in the progress of $mser_{rel}$ and $mser_{abs}$. Since this is (from 1b) exactly the range, where using single- or recompressed data does have an impact, we propose - based on observations 1d and 1c - to bench-mark compression algorithms in respect to iris segmentation by using

- uncompressed data sets rated with relative measures, such as the $mser_{rel}$, for severe compression, i.e. $cr_t > 50$ and
- pre-compressed data sets[5] with absolute measures, such as $mser_{abs}$, for medium and light compression, i.e. $cr_t \leq 50$.

We were not able to determine, whether or in which way the choice of the pre-compression parameter relates to the level of segmentation error rates (1e).

In conclusion, one has to be careful when using pre-compressed data for compression bench-marking, when the pre-compression as well as the bench-marked compression method were the same, i.e. JPEG. If this applies to intra-compression with other methods as well needs further investigation.

In inter-recompression we cannot observe such behaviour. However, there are trends observeable (2a, 2b), which need further investigation.

5. Comparison

Besides evaluating the segmentation error rate and the general purpose quality measures independently, their correlation is analysed. Furthermore, we explicitly investigate the correlation between $mser_{rel}$ and $mser_{abs}$ to back up the observations in section 4.2. For this purpose, we used the Spearman's rank correlation coefficient (SRCC).

5.1. Correlation of Evaluation methods

5.1.1 Quality Metrics vs. Segmentation Error Rate

In general, all five used general purpose quality metrics show a high linear relationship with a minimum SRCC of 0.852 to each other. Furthermore, all of these metrics are highly correlating with the segmentation error rate and there are only minor differences between the segmentation algorithms (WAHET and CAHT) and the relative and absolute segmentation error rates.

Table 1 shows the correlation results for the intra-compression (*jpg*) and the CAHT segmentation algorithm. It can be observed that the $mser_{rel}$ shows overall a higher linear relationsship to the quality metrics than the $mser_{abs}$. The reason for this can be seen in Figure 6 (bottom) where the $mser_{rel}$ has a higher slope for small and medium CRs where the $mser_{rel}$ is more flat in this region and is in general more noisy as well. Furthermore, it can be seen that for the MSSIM metric ($mser_{rel}$) in case of single compression the SRCC is smaller compared to the other metrics, which is due to the outlier of the MSSIM at $cr_t = 15$ in Figure 3. However, except of this outlier and when looking at the graphs as well, the MSSIM represents best the segmentation error rates in case of the intra-compression, because the single compressed curve gets away from the bundle of the recompressed curves. On the other hand, on jp2 compressed data (figure 3) MSSIM shows absolutely no sign of the stagnancy, although there is clear stagnancy in figure 7 (left). Hence, MSSIM (among others) cannot be used to predict behaviour of iris segmentation. This could also mean that the segmentation not only depends on the quality of an image, or at least not in a linear relationship.

SRCC between the quality metrics and the relative MSER					
	single	*rec.70*	*rec.75*	*rec.80*	*rec.100*
PSNR	-0.995	-0.995	-1.0	-1.0	-1.0
MSSIM	-0.912	-0.995	-1.0	-1.0	-1.0
NQM	-0.995	-0.995	-1.0	-1.0	-1.0
VSNR	-0.995	-0.995	-1.0	-1.0	-1.0
RFSI	-0.995	-0.995	-1.0	-1.0	-0.995

SRCC between the quality metrics and the absolute MSER					
	single	*rec.70*	*rec.75*	*rec.80*	*rec.100*
PSNR	-0.857	-0.863	-0.929	-0.868	-0.841
MSSIM	-0.923	-0.863	-0.929	-0.868	-0.841
NQM	-0.857	-0.863	-0.929	-0.868	-0.841
VSNR	-0.857	-0.863	-0.929	-0.868	-0.841
RFSI	-0.857	-0.863	-0.929	-0.868	-0.846

Table 1. Spearman Rank Correlation Coefficient between the quality metrics and the $mser_{rel}$ (above) as well as the $mser_{abs}$ (below) for intra-compression (*jpg*) and the CAHT segmentation algorithm.

5.1.2 Relative vs. Absolute Segmentation Error Rate

The SRCC between the relative and absolute segmentation error rates confirms that both metrics show in general the same trend as it can be seen in Table 2.

In case of the WAHET segmentation *jp2* outperforms the other two compression methods. The smaller SRCC values for *jpg* is mainly due to the $mser_{abs}$ at smaller compression

[5] If absolutely necessary because of ground-truth availability, of course uncompressed data is preferred

WAHET Segmentation					
	single	*rec.70*	*rec.75*	*rec.80*	*rec.100*
jpg	0.703	0.835	0.890	0.786	0.863
j2k	0.978	0.962	0.978	0.923	0.978
jxr	0.742	0.956	0.423	0.544	0.412

CAHT Segmentation					
	single	*rec.70*	*rec.75*	*rec.80*	*rec.100*
jpg	0.863	0.802	0.928	0.868	0.841
j2k	0.984	1.0	0.995	1.0	1.0
jxr	0.978	0.973	0.967	0.918	0.978

Table 2. SRCC between $mser_{rel}$ and $mser_{abs}$ for all three methods and both WAHET segmentation (above) and CAHT segmentation (below).

ratios where the segmentation error rate is higher for $cr_t = 15$ than for $cr_t = 20$. Also the small ripple at $cr_t = 30$ has an impact on the correlation here. The reason for the intermediate SRCC results in case of jxr are mainly due to noise, which can be seen in Figure 7.

6. Conclusion

References

[1] H. Bauschke, C. Hamilton, M. Macklem, J. S. McMichael, and N. Steward. Recompression of jpeg images by requantization. *IEEE Transactions on Image Processing*, 12(7):843–849, 2003.

[2] K. W. Bowyer and P. J. Flynn. The nd-iris-0405 iris image dataset. Technical report.

[3] M. Burge and K. Bowyer, editors. *Handbook of Iris Recognition*. 2013.

[4] S. Chan. Recompression of still images. Technical Report 21068, Computing Laboratory, University of Kent, UK, 1992.

[5] D. M. Chandler and S. Hemami. Vsnr: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, vol. 16, no. 9:pp. 2284–2298, 2007.

[6] J. Daugman and C. Downing. Effect of severe image compression on iris recognition performance. *IEEE Transactions on Information Forensics and Security*, 3, 2008.

[7] J. Hämmerle-Uhl, C. Prähauser, T. Starzacher, and A. Uhl. Improving compressed iris recognition accuracy using JPEG2000 RoI coding. In M. Tistarelli and M. Nixon, editors, *Proceedings of the 3rd International Conference on Biometrics 2009 (ICB'09)*, volume 5558 of *LNCS*, pages 1102–1111. Springer Verlag, 2009.

[8] H. Hofbauer, F. Alonso-Fernandez, P. Wild, J. Bigun, and A. Uhl. A ground truth for iris segmentation. In *Proceedings of the 22th International Conference on Pattern Recognition (ICPR'14)*, page 6pp., Stockholm, Sweden, 2014.

[9] K. Horvath, H. Stögner, and A. Uhl. Effects of JPEG XR compression settings on iris recognition systems. In P. Real, D. Diaz-Pernil, H. Molina-Abril, A. Berciano, and W. Kropatsch, editors, *Proceedings of the 14th International Conference on Computer Analysis of Images and Patterns (CAIP 2011)*, volume 6855 of *LNCS*, pages 73–80. Springer Verlag, 2011.

[10] I.O. for Standardization. ISO/IEC IS 10918-1. Digital compression and coding of continuous-tone still images: Requirements and guidelines, 1994.

[11] I.O. for Standardization. ISO/IEC IS 15444-1. Information technology JPEG 2000 image coding system: Core coding system, 2004.

[12] I.O. for Standardization. ISO/IEC IS 29199-2. Information technology JPEG XR image coding system Part 2: Image coding speci?cation, 2012.

[13] R. W. Ives, D. Bishop, Y. Du, and C. Belcher. Iris recognition: The consequences of image compression. *EURASIP Journal of Advances in Signal Processing*, 2010:Article ID 680845, doi:10.1155/2010/680845, 2010.

[14] R. W. Ives, R. P. Broussard, L. R. Kennell, and D. L. Soldan. Effects of image compression on iris recognition system performance. *Journal of Electronic Imaging*, 17:011015, doi:10.1117/1.2891313, 2008.

[15] M. Konrad, H. Stögner, and A. Uhl. Custom design of JPEG quantization tables for compressing iris polar images to improve recognition accuracy. In M. Tistarelli and M. Nixon, editors, *Proceedings of the 3rd International Conference on Biometrics 2009 (ICB'09)*, volume 5558 of *LNCS*, pages 1091–1101. Springer Verlag, 2009.

[16] X. M. L. Zhang, L. Zhang. Rfsim: A feature based image quality asesment metric using riesz transforms. *Proc. 17th IEEE Int. Conf. on Image Procesing ICIP*, page p. 321324, 2010. Hong Kong, China.

[17] S. Matschitsch, M. Tschinder, and A. Uhl. Comparison of compression algorithms' impact on iris recognition accuracy. In S.-W. Lee and S. Li, editors, *Proceedings of the 2nd International Conference on Biometrics 2007 (ICB'07)*, volume 4642 of *LNCS*, pages 232–241. Springer Verlag, 2007.

[18] W. S. G. B. L. E. N. Damera-Venkata, T. D. Kite and A. C. Bovik. Image quality assessment on a degradation model. *IEEE Transactions on Image Processing, vol. 9, no. 4, pp. 636-649*, 2000.

[19] B. Pardamean and I. Christiani. Comparatative evaluation of iris recognition performance. *International Journal of Mathematical Models and Methods in Applied Sciences*, 6(2):332–339, 2012.

[20] W. Pennebaker and J. Mitchell. *JPEG – Still image compression standard*. Van Nostrand Reinhold, New York, 1993.

[21] P. Philips, K. Bowyer, and P. Flynn. Comments on the CASIA version 1.0 iris data set. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(10):1869–1870, Oct. 2007.

[22] S. Rakshit and D. Monro. An evaluation of image sampling and compression for human iris recognition. *IEEE Transactions on Information Forensics and Security*, 2(3):605–612, 2007.

[23] C. Rathgeb, A. Uhl, and P. Wild. *Iris Recognition: From Segmentation to Template Security*. Springer, New York, 2012.

[24] C. Rathgeb, A. Uhl, and P. Wild. *Iris Recognition: From Segmentation to Template Security*, volume 59 of *Advances in Information Security*. Springer Verlag, 2013.

[25] C. Rathgeb, A. Uhl, and P. Wild. Effects of severe image compression on iris segmentation performance. In *IJCB 2014*, September 2014.

[26] H. Sencar and N. M. (Eds.). *Digital Image Forensics: There is more to a picture than meets the eye*. Springer Verlag, 2012.

[27] A. Tuba, J. Arezina, and R. Jovanovic. Software framework for testing JPEG quantization tables for iris recognition. In *Applied Mathematics in Electrical and Computer Engineering*, pages 417–422, 2012.

[28] E. P. S. Z. Wang and A. C. Bovik. Multi-scale structural similarity for image quality assessment. *IEEE Asiloma Conference on Signal, Systems and Computers*, 2003.