Optimisation of JPEG XR Quantisation Settings in Iris Recognition Systems

Kurt Horvath and Herbert Stögner School of Communication Engineering for IT Carinthia University of Applied Sciences Klagenfurt, Austria

KurtKlaus.Horvath@edu.fh-kaernten.ac.at

Abstract—JPEG XR is considered as a lossy sample data compression scheme in the context of iris recognition techniques. It is shown that by optimising the JPEG XR quantisation strategy, JPEG XR default quantisation as well as JPEG2000 based iris recognition can be improved in terms of EER. The optimised JPEG XR quantisation strategy shows good performance across a wide range of iris feature extraction techniques, but has to be adapted for each target bitrate separately.

Keywords-JPEG XR; iris recognition; quantisation optimisation; EER.

I. INTRODUCTION

In distributed biometric systems, the compression of sample data may become imperative under certain circumstances, since the data acquisition stage is often dislocated from the feature extraction and matching stage. In such environments the sample data have to be transferred via a network link to the respective location, often over wireless channels with low bandwidth and high latency. Therefore, a minimisation of the amount of data to be transferred is highly desirable, which is achieved by compressing the data before transmission and any further processing. See Fig. 1 for an illustration involving JPEG XR for compressed data transmission.



Fig. 1. System View.

As an alternative, the application of feature extraction before transmission looks promising due to the small size of template data but cannot be done under most circumstances due to the prohibitive computational demand of these operations (current sensor devices are typically far too weak to support this Andreas Uhl Multimedia Signal Processing and Security Lab Department of Computer Sciences University of Salzburg, Austria andreas.uhl@sbg.ac.at

while compression can be done e.g. in dedicated low power hardware).

In order to maximise the benefit in terms of data reduction, lossy compression techniques are often suggested. Given the potential impact of lossy compression techniques on biometric recognition performance, it is imperative to carefully select and optimise appropriate codecs and to study their corresponding effect on recognition accuracy.

While current international standards define the application of JPEG2000 for lossy iris sample data compression, we focus in this paper on the optimised application of the recent JPEG XR still image coding standard. We experimentally compare the achieved results to a JPEG2000 based (and therefore standard conformant) environment. In particular, besides reviewing the effects of applying different settings concerning the use of the optional Photo Overlap Transform (POT) as a part of JPEG XR's Lapped Biorthogonal Transform (LBT), we optimise the JPEG XR quantisation strategy with respect to balancing quantisation strength among the three different frequency bands of the LBT. In Section 2, we review related standards and literature in the area of lossy iris sample data compression, while in Section 3, JPEG XR basics and especially the quantisation strategy are briefly explained. Section 4 presents experiments where we first shortly review the four different iris recognition systems employed in this study. Subsequently, the optimisation of the JPEG XR quantisation scheme is explained. Experimental results comparing optimised JPEG XR, different LBT variants in JPEG XR, and JPEG2000 are shown with respect to iris recognition accuracy in terms of EER. Section 5 concludes the paper.

II. BIOMETRIC IRIS SAMPLE COMPRESSION

During the last decade, several algorithms and standards for compressing image data relevant in biometric systems have evolved. The certainly most relevant one is the ISO/IEC 19794 standard on Biometric Data Interchange Formats, where in its former version (ISO/IEC 19794-6:2005), JPEG and JPEG2000 (and WSQ for fingerprints) were defined as admissible formats for lossy compression, whereas for lossless and nearly lossless compression JPEG-LS as defined in ISO/IEC 14495 was suggested. In the most recently published version (ISO/IEC FDIS 19794-6 as of August 2010), only JPEG2000 is included for lossy compression while the PNG format serves as lossless compressor [1]. These formats have also been recommended for various application scenarios and standardised iris images (IREX records) by the NIST Iris Exchange¹ program.

The ANSI/NIST-ITL 1-2011 standard on "Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information" (2nd draft as of February 2011, former ANSI/NIST-ITL 1-2007) supports both PNG and JPEG2000 for the lossless case and JPEG2000 only for applications tolerating lossy compression.

In literature on compressing iris imagery, rectilinear [2], [3], [4], [5] as well as polar [6], [7], [8], [9] iris sample data formats has been considered. With respect to employed compression technology, we find JPEG [3], [4], [5], JPEG2000 [2], [3], [4], [5], and other general purpose compression techniques [4], [5] being investigated. Superior compression performance of JPEG2000 over JPEG is seen especially for low bitrates (thus confirming the choice of the above-referenced standards), however, for high and medium quality, JPEG is found still to be competitive in terms of impacting recognition accuracy. Apart from applying the respective algorithms with their default settings and standard configurations, work has been done to optimise the compression algorithms to the application domain: For JPEG2000, it has been proposed to invoke RoI coding for the iris texture area [10] whereas the removal of the image background before compression has also been suggested (i.e. parts of the image not being part of the eye like eyelids are replaced by constant average gray [3]). For JPEG, an optimisation of quantisation matrices has been proposed to achieve better matching accuracy compared to the standard values for rectangular iris image data [11] as well as for polar iris images [8], [9].

The JPEG XR standard has only recently been investigated in the context of biometric systems [12]. It has been found to eventually represent an interesting alternative to JPEG2000 in iris recognition systems due to its simpler structure and less demanding implementations in terms of memory and CPU resources, while providing almost equal recognition performance.

III. JPEG XR BACKGROUND

Originally developed by Microsoft and termed "HD Photo", JPEG XR got standardised by ITU-T and ISO in 2009 [13], which makes it the most recent still image coding standard. The original scope was to develop a coding scheme targeting "extended range" applications which involves higher bitdepths as currently supported. However, much more than 10 years after JPEG2000 [14] development and 10 years after its standardisation it seems to be reasonable to look for a new coding standard to eventually employ "lessons learnt" in JPEG2000 standardisation. In particular, the focus is on a simpler scheme which should offer only the amount of scalability actually required for most applications (as opposed

¹http://iris.nist.gov/irex/

to JPEG2000 which is a rather complex scheme offering almost unconstraint scalability).

JPEG XR is a transform coding scheme showing the classical three-stage design: transform, quantisation, and entropy encoding. The transform operates on macroblocks consisting of 16 (arranged in 4 by 4) 4×4 pixel blocks. The first stage of the integer-based transform is applied to all 4×4 pixel blocks of a macroblock. Subsequently, the resulting coefficients are partitioned into 240 "high pass (HP) coefficients" and 16 coefficients corresponding to the lowest frequency in each block. The latter are aggregated into a square data layout (4 x 4 coefficients) onto which the transform is applied for a second time. The result are 15 "low pass (LP) coefficients" and a single "DC" coefficient (per macroblock).

In fact, the transform used in JPEG XR is more complicated as compared to JPEG, it is a so-called "two-stage lapped biorthogonal transform (LBT)" which is actually composed of two distinct transforms: The Photo Core Transform (PCT) and the Photo Overlap Transform (POT). The PCT is similar to the widely used DCT and exploits spatial correlation within the 4 x 4 pixels block, however, it suffers from the inability to exploit inter-block correlations due to its small support and from blocking artifacts at low bitrates. The POT is designed to exploit correlations across block boundaries as well as to mitigate blocking artifacts.



Fig. 2. Rectilinear example images.

Each stage of the transform can be viewed as a flexible concatenation of POT and PCT since the POT is functionally independent of the PCT and can be switched on or off, as chosen by the encoder (this is signalled by the encoder in the bitstream). There are three options: disabled for both PCT stages (LBT=0), enabled for the first PCT stage but disabled for the second PCT stage (LBT=1), or enabled for both PCT stages (LBT=2). In recent work is has been shown

that surprisingly, no clear advantage of any of these options with respect to recognition performance can be observed [12].



Fig. 3. No compression applied.

Fig. 2 shows sample images for the uncompressed case and the three transform settings of JPEG XR (LBT=0,1,2) with "uniform" quantisation parameter q = 100 (see below).



Fig. 4. LBT=0, HD=0.35.

Figs. 3 - 6 visualise corresponding extracted iris textures as well as computed Masek Iris Codes (see next section) for the four settings shown in Fig. 2. When computing the Hamming Distance (HD) to the iris code derived from the uncompressed image in Fig. 3, we result in 0.35 for LBT=0, 0.403 for LBT=1, and 0.385 for LBT=2.



Fig. 5. LBT=1, HD=0.403.

In this work we specifically focus on the quantisation strategy in JPEG XR. After the LBT transform, the coefficients in the DC,LP,HP bands are quantised by a (integer) value q in the range 1 - 255. In the case of "uniform" quantisation (which is the default setting), all three bands are quantised with the same value. For controlling the amount of compression, q is scaled but can only be of integer type. However, JPEG XR also allows to apply different quantisation parameters for the DC, LP, and HP subbands besides the uniform strategy (in any case, the coefficients within one of these subbands are all quantised with an identical value). This corresponds to giving different emphasis to low frequency (DC band), mid frequency (LP band), and high frequency (HP band) information, respectively.

The aim of this work is to optimise the quantisation parameter settings for the three DC,LP,HP bands in the context of iris recognition instead of applying the default uniform strategy. Results will also shed light on the question which



(b) Iris Code

Fig. 6. LBT=2, HD=0.385.

frequency bands do carry the most discriminative information in iris imagery.

Since our experiments are focused on the evaluation of those quantisation-related questions, we do not describe the subsequent JPEG XR stages in the following, please consult the standard or related publications with respect to these issues [13].

IV. EXPERIMENTS ON OPTIMISING JPEG XR COMPRESSION OF IRIS SAMPLE DATA

A. Iris Recognition and Iris Database

It is crucial to assess the effects of compressing iris samples using a set of different iris recognition schemes since it can be expected that different feature extraction strategies will react differently when being confronted with compression artefacts and reduced image quality in general.

Many iris recognition methods follow a quite common scheme [15], close to the well known and commercially most successful approach by Daugman [16]. In our pre-processing approach (following e.g. Ma et al. [17]) we assume the texture to be the area between the two almost concentric circles of the pupil and the outer iris. These two circles are found by contrast adjustment, followed by Canny edge detection and Hough transformation. After the circles are detected, unwrapping along polar coordinates is done to obtain a rectangular texture of the iris. In our case, we always re-sample the texture to a size of 512x64 pixels. Subsequently, features are extracted from this iris texture (which has also been termed polar iris image). We consider the following four techniques in this work, which are selected to represent a broad variety of different template generation concepts:

(1) A wavelet-based approach proposed by Ma et al. [17] is used to extract a bit-code. The texture is divided into N stripes to obtain N one-dimensional signals, each one averaged from the pixels of M adjacent rows. We used N = 10 and M = 5for our 512x64 pixel textures (only the 50 rows close to the pupil are used from the 64 rows, as suggested in [17]). A dyadic wavelet transform is then performed on each of the resulting 10 signals, and two fixed subbands are selected from each transform. This leads to a total of 20 subbands. In each subband we then locate all local minima and maxima above some threshold, and write a bitcode alternating between 0 and 1 at each extreme point. Using 512 bits per signal, the final code is then 512x20 bit. Matching different codes is done by computing the Hamming Distance.

(2) Again restricting the texture to the same N = 10 stripes as described before, we use a custom C implementation similar

to Libor Masek's Matlab implementation² of a 1-D version of the Daugman iris recognition algorithm as the second feature extraction technique. A row-wise convolution with a complex Log-Gabor filter is performed on the texture pixels. The phase angle of the resulting complex value for each pixel is discretized into 2 bits. Those 2 bits of phase information are used to generate a binary code, which therefore is 512x20bit (again, Hamming Distance can be used for similarity determination).

(3) The third algorithm has been proposed by Ko et al. [18]. Here feature extraction is performed by applying cumulativesum-based change analysis. The algorithm discards parts of the iris texture, from the right side $[45^{\circ} \text{ to } 315^{\circ}]$ and the left side $[135^{\circ} \text{ to } 225^{\circ}]$, since the top and bottom of the iris are often hidden by eyelashes or eyelids. Subsequently, the resulting texture is divided into basic cell regions (these cell regions are of size 8×3 pixels). For each basic cell region an average gray scale value is calculated. Then basic cell regions are grouped horizontally and vertically (one group consists of five basic cell regions). Finally, cumulative sums over each group are calculated to generate an iris-code. If cumulative sums are on an upward slope or on a downward slope these are encoded with 1s and 2s, respectively, otherwise 0s are assigned to the code. In order to obtain a binary feature vector (to enable Hamming Distance computation for comparison) we rearrange the resulting Iris Code such that the first half contains all upward slopes and the second half contains all downward slopes. With respect to the above settings the final iris-code consists of 2400 bits.

(4) Finally, we employ the feature extraction algorithm of Zhu et al. [19] which applies a 2-D wavelet transform to the polar image first. Subsequently, first order statistical measures are computed from the wavelet subbands (i.e. mean and variance) and are concatenated into a feature vector. The similarity between two of these real-valued feature vectors is determined by computing the corresponding l^2 -Norm.

We used the CASIAv3 Interval dataset³ in the experiments. It consists of NIR images with 320×280 pixels in 8 bit grayscale .jpeg format (high quality) of 249 persons, where for many persons both eyes are available which leads to 391 (image) classes overall.

For intra-class matches (genuine user matches), we consider all possible template pairs for each class (overall 8882 matches), while for inter-class matches (impostor matches) the first two templates of the first person are matched against all templates of the other classes (overall 2601 matches).

B. Compression Techniques Settings

In JPEG XR quantisation, we aim at optimising the relation among the quantisation parameters for the three subbands DC, LP, and HP, i.e. we look for the triple q:r:s which provides the best solution in terms of recognition performance (measured in equal error rate (EER)). Since it is not obvious that there

²http://www.csse.uwa.edu.au/~pk/studentprojects/ libor/sourcecode.html exists a unique optimal solution independent of target bit rate, we look for an optimal q:r:s triple with respect to a certain target bitrate. Since the number of q:r:s triples is way too large to be tested exhaustively, we have quantised the search space into 18 DC bands, and 15 LP and 15 HP bands, respectively. Still 4050 possible combinations need to be considered, but this is more tractable compared to $255^3 = 16581375$ triples without quantisation.

For enabling a fair comparison between the various quantised triples in the experiments, the same bitrate has to be targeted for all configurations. While specifying a target bitrate is straightforward in JPEG2000, JPEG XR suffers from the same weakness as JPEG being unable to explicitly specify a target bitrate. Therefore we have employed a wrapperprogram, continuously scaling the JPEG XR quantisation triples (i.e. multiplication of all three components with the same factor) to achieve a certain target bitrate (given in bytes per pixel bpp). Since q,r,s can attain integer values only, target bitrates are approximated as accurate as possible. In Fig. 7 we show an example of approximating a target bitrate of 0.1968 bytes/pixel for more then 2500 images. On average we get 0.1966 bytes/pixel with a maximal deviation of +5.97% and -6.21%.



Fig. 7. Rate adaptation approximation.

For experimentation, we use the official JPEG-XR reference software 1.8 (as of September 2009) and for JPEG2000 compression, imagemagick 8.6.6.0.4-3 (employing libJASPER 1.900.1-7+b1) is used with standard settings.

The optimisation is done minimising the EER of the Masek implementation by setting LBT=0 since this is the fastest variant and there are no clear recognition advantages of using LBT=1,2 [12].

The questions we want to answer with our experiments are as follows:

- 1) Do the optimised settings outperform the "uniform" JPEG XR default settings ?
- 2) Do the optimised settings outperform JPEG2000 ?
- 3) Do the optimised settings also generalise to other bitrates (since they have been computed for a single target bitrate) ?
- 4) Do the optimised settings also generalise to other feature extraction schemes (since they have been computed for

³http://www.cbsr.ia.ac.cn/IrisDatabase.htm/

the Masek Iris Code) ?

C. Experimental Results

Fig. 8 shows computed tuples r:s, when all triples are normalised with q = 1. Out of all considered 4050 r:s(:1) triples, blue dots show configurations when the obtained EER is at least 5% better as compared to uniform q:r:s, and red diamonds depict configurations with at least 15% improvement. The target bitrate for the optimisation has been set to 0.19 bytes/pixel (filesize is 17 kBytes) for all experimental results shown. Note that experiments with different target bitrates lead to highly similar results with respect to the answers to the four questions raised above, but of course not with respect to the actual triples q:r:s computed.



Fig. 8. Result Distribution

We clearly note that the best triples are not close to the uniform setting q:r:s = 1 but 1 < r < 2 and 2.9 < s < 4. This means that the higher frequency gets, the more severe quantisation should be applied.

Fig. 9 shows the results of two good q:r:s configurations for varying the bitrate in compression (x-axis) and performing iris recognition with the Masek Iris Code EER is plotted on the y-axis). For a comparison, we plot the curves for LBT=0,1,2 with uniform q:r:s and a curve obtained from applying JPEG2000. For both configurations we observe that for the optimisation target bitrate, the optimised q:r:s triple is clearly superior to the "uniform" JPEG XR variants and also superior to JPEG2000.

However, this superiority does not at all extend to other bitrates. The bitrate range where these triples exhibit better performance is quite limited. This means that in an application, specific q:r:s triples need to be optimised for different target bitrates. The behaviour of those two configurations as shown in Figs. 9.a and 9.b is very similar except for the the range of bytes/pixel < 0.15. Here the better preservation of LP and to a lesser extent HP data for q:r:s = 1:1.19:2.93 leads to performance close or even better to JPEG2000 (see Fig. 9.a). Note that for bitrates > 0.05, in many cases EER derived from lossy compression is superior to the values computed from uncompressed data - this effect has been observed in many studies and is due to the de-noising effect of moderate compression settings.



Fig. 9. Recognition with Masek Iris Code.

Finally, we want to answer the question in how far the good results of the computed triples do generalise to different types of feature extraction schemes and resulting Iris Codes without explicit optimisation for the respective algorithms.

In Fig. 10, we compare the behaviour of the three remaining feature extraction techniques when applied to sample data which have been compressed using the triple q:r:s = 1:1.97:3.15 – which has been optimised for the Masek Iris Code at bitrate 0.19 bytes/pixel. We notice that for the target bitrate, the EER values are fairly good for all three types of iris codes. While for the Ma and Ko variants, the result is better compared to JPEG2000 and all three uniform variants, the Zhu variant is slightly inferior to LBT=2 only, but superior to all other compression schemes including JPEG2000. So it seems that this q:r:s configuration is able to preserve texture information for the targeted bitrate very well, no matter which subsequent feature extraction technique is being applied.

On the other hand, we notice again that the bitrate range where this good behaviour is observed is actually quite limited (except for the Ko Iris Code, where we see good results for lower bitrates also). The specifically good results at bitrate 0.05 bytes/pixel for the Ko and Zhu feature extraction schemes are probably due to optimal denoising behaviour at this compression ratio for these two schemes.

V. CONCLUSION

We have found that optimising the JPEG XR quantisation strategy leads to improved iris recognition results for a wide



range of different feature extraction types. The optimised strategy does not only outperform the default quantisation strategy but also iris recognition relying on JPEG2000 compression. The observed behaviour is only found in a small range of bitrates close to the target bitrate that has been used for optimisation, however, the optimised parameters for a specific feature extraction technique do also provide good results for other types of Iris Codes. The general trend with respect to the importance of different frequency bands is that as opposed to the JPEG XR default configuration, middle LP frequencies and even more pronounced high HP frequencies should be quantised more severely compared to the low frequency DC information.

REFERENCES

[1] K. Horvath, H. Stögner, A. Uhl, and G. Weinhandel, "Lossless compression of polar iris image data," in *Proceedings of the 5th Iberian* *Conference on Pattern Recognition and Image Analysis (IbPRIA 2011)*, ser. LNCS, J. Vitria, J. M. Sanches, and M. Hernandez, Eds., vol. 6669. Springer Verlag, 2011, pp. 329–337.

- [2] 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*, vol. 17, pp. 011 015, doi:10.1117/1.2 891 313, 2008.
- [3] J. Daugman and C. Downing, "Effect of severe image compression on iris recognition performance," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 52–61, 2008.
- [4] S. Matschitsch, M. Tschinder, and A. Ühl, "Comparison of compression algorithms' impact on iris recognition accuracy," in *Proceedings of the* 2nd International Conference on Biometrics 2007 (ICB'07), ser. LNCS, S.-W. Lee and S. Li, Eds., vol. 4642. Springer Verlag, 2007, pp. 232– 241.
- [5] S. Jenisch, S. Lukesch, and A. Uhl, "Comparison of compression algorithms' impact on iris recognition accuracy II: revisiting JPEG," in *Proceedings of SPIE, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, vol. 6819, San Jose, CA, USA, Jan. 2008, p. 68190M ff.
- [6] S. Rakshit and D. Monro, "Effects of sampling and compression on human iris verification," in *Proceedings of the IEEE International Conference on Acustics, Speech, and Signal Processing (ICASSP 2006)*, Tolouse, France, 2006, pp. II–337–II–340.
- [7] —, "An evaluation of image sampling and compression for human iris recognition," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 605–612, 2007.
- [8] M. Konrad, H. Stögner, and A. Uhl, "Custom design of JPEG quantization tables for compressing iris polar images to improve recognition accuracy," in *Proceedings of the 3rd International Conference on Biometrics 2009 (ICB'09)*, ser. LNCS, M. Tistarelli and M. Nixon, Eds., vol. 5558. Springer Verlag, 2009, pp. 1091–1101.
- [9] —, "Evolutionary optimization of JPEG quantization tables for compressing iris polar images in iris recognition systems," in *Proceedings of* the 6th International Symposium on Image and Signal Processing and Analysis, ISPA '09, Salzburg, Austria, Sep. 2009.
- [10] J. Hämmerle-Uhl, C. Prähauser, T. Starzacher, and A. Uhl, "Improving compressed iris recognition accuracy using JPEG2000 RoI coding," in *Proceedings of the 3rd International Conference on Biometrics 2009* (*ICB'09*), ser. LNCS, M. Tistarelli and M. Nixon, Eds., vol. 5558. Springer Verlag, 2009, pp. 1102–1111.
- [11] G. Kostmajer, H. Stögner, and A. Uhl, "Custom JPEG quantization for improved iris recognition accuracy," in *Emerging Challenges for Security, Privacy and Trust. Proceedings of the 24th IFIP International Information Security Conference 2009 (IFIP SEC'09)*, ser. IFIP AICT, D. Gritzalis and J. Lopez, Eds., vol. 297. Springer Verlag, May 2009, pp. 76–86.
- [12] K. Horvath, H. Stögner, and A. Uhl, "Effects of JPEG XR compression settings on iris recognition systems," in *Proceedings of the 14th International Conference on Computer Analysis of Images and Patterns (CAIP* 2011), ser. LNCS, P. Real, D. Diaz-Pernil, H. Molina-Abril, A. Berciano, and W. Kropatsch, Eds., vol. 6855. Springer Verlag, 2011, pp. 73–80.
- [13] F. Dufaux, G. J. Sullivan, and T. Ebrahimi, "The JPEG XR image coding standard," *IEEE Signal Processing Magazine*, vol. 26, no. 6, pp. 195– 199, Nov. 2009.
- [14] D. Taubman and M. Marcellin, JPEG2000 Image Compression Fundamentals, Standards and Practice. Kluwer Academic Publishers, 2002.
- [15] K. Bowyer, K. Hollingsworth, and P. Flinn, "Image understanding for iris biometrics: A survey," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281 – 307, 2008.
- [16] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.
- [17] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Efficient iris recognition by characterizing key local variations," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 739–750, Jun. 2004.
- Processing, vol. 13, no. 6, pp. 739–750, Jun. 2004.
 [18] J.-G. Ko, Y.-H. Gil, J.-H. Yoo, and K.-I. Chung, "A novel and efficient feature extraction method for iris recognition," *ETRI Journal*, vol. 29, no. 3, pp. 399 401, 2007.
- [19] Y. Zhu, T. Tan, and Y. Wang, "Biometric personal identification based on iris patterns," in *Proceedings of the 15th International Conference on Pattern Recognition (ICPR'00)*, vol. 2. IEEE Computer Society, 2000, pp. 2801–2804.