Highly Efficient Protection of Biometric Face Samples with Selective JPEG2000 Encryption

Heinz Hofbauer[1] • Yoanna Martínez-Díaz[2] • Simon Kirchgasser[1] • Heydi Méndez-Vázquez[2] • Andreas Uhl[1]

[1]Multimedia Signal Processing and Security Lab, Paris Lodron University of Salzburg, Austria
{hofbauer, skirch, uhl}@cs.sbg.ac.at
[2] Advanced Technologies Application Center (CENATAV), Havana, Cuba
{ymartinez, hmendez}@cenatav.co.cu

February 9, 2021

Abstract

When biometric databases grow larger, a security breach or leak can affect millions. In order to protect against such a threat, the use of encryption is a natural choice. However, a biometric identification attempt then requires the decryption of a potential huge database, making a traditional approach potentially unfeasible. The use of selective JPEG2000 encryption can reduce the encryption's computational load and enable a secure storage of biometric sample data. In this paper we will show that selective encryption of face biometric samples is secure. We analyze various encoding settings of JPEG2000, selective encryption parameters on the "Labeled Faces in the Wild" database and apply several traditional and deep learning based face recognition methods.

Contents

1 Introduction

The International Organisation for Standardisation (ISO) specifies biometric data to be recorded and stored in (raw) image form (ISO/IEC FDIS 19794), i.e., sample images, and not only in extracted templates. Such deployments benefit from future improvements which can be incorporated without re-enrollment of registered users, thereby increasing interoperability and vendor neutrality [1]. In ISO/IEC Standard 19794-5:2011 [2] two encodings are defined for facial images: JPEG and JPEG2000. Several studies have been performed, e.g. [3] and [4], recommending to use JPEG2000 compression instead JPEG. Therefore, we will also use JPEG 2000 in this study.

In a biometric system the biometric templates are stored in an online-database, optimally protected via template protection (TP) schemes. The corresponding biometric sample data will certainly be stored off-line, optimally protected via encryption by a state of the art cipher, e.g. AES. Whenever a sample from the offline database needs to be accessed it needs to be *decrypted* first, which is time consuming, in particular when the whole database needs to be accessed. This is the case when (1) the biometric comparison or template extraction technique is changed (the alternative would be a re-enrollment of all users); and (2) the key used in the employed TP scheme needs to be changed due to a periodic update (as a preventive measure to guard against undetected loss of a key) or because it has been lost in an attack or data breach (to reestablish security). Other scenarios are given if one or several samples in the offline database need to be decrypted include: (1) template regeneration for single users; (2) explicit sample comparison in forensic identification where a human operator typically confirms automatically pre-selected biometric correspondences; or (3) de-duplication, where, similar to the latter case, sample data is compared for de-duplication purposes. Another entirely different application scenario for the time-critical protection of sample data is in distributed biometric recognition. Here biometric sample data is sent from an acquisition device to the authentication component and can be intercepted by an eavesdropper on the channel. As sensors typically do not exhibit high computational performance and JPEG2000 compression can be conduced in hardware, e.g., by a dedicated chip, the subsequent *encryption* of the sample before transmission needs to be low cost. We investigate a lightweight JPEG2000 encryption scheme for compressed face data, based on selective bit-stream protection using AES. The proposed technique offers the benefits of traditional encryption, i.e., security, no reduced recognition accuracy and combines them with a low computational effort. This comes at the drawback that a certain amount of data is left in plain text and could be used for an attack against the system. It is required to perform a security analysis similar to those done for template protection (as the same attack scenarios apply). In particular, we have to consider the interplay between different JPEG2000 encoding types, layer and resolution progression, and the data selection for encryption (amount and position). The amount of encryption that is required gives the speedup (the benefit) over traditional encryption, e.g., when half the data has to be encrypted the resulting speedup would be two.

Section 2 contains a brief overview of current literature and introduces our approach (Section 2.2). Section 3 shows the results of the experimental evaluation and discusses relevant findings. Finally, Section 4 summarizes all findings from the experimental evaluation.

2 Related Work and Evaluation Methodology

2.1 Selective JPEG2000 Encryption Methods

A large variety of custom image encryption schemes have been developed for JPEG2000 [5], many of them being motivated by the potential reduction of computational effort as compared to full encryption. Reducing computational encryption effort is of interest in the context of biometric systems in case either weak hardware (e.g. mobile sensing devices) or large quantities of data (e.g. nation-wide sample databases) are involved.

To enable security assessment (which involves decoding of encrypted data), only format compliant encryption schemes are admissible. Thus, we apply a format compliant JPEG2000 encryption scheme introduced in the context of JPSEC [6] to avoid such pitfalls.

When assessing the security of format compliantly encrypted visual data, the data can simply be decoded with the encrypted parts. Due to format compliance, this is possible with any given decoding scheme, however, the encrypted parts introduce noise-type distortion to the data. A highly efficient attack is replacing the encrypted parts with carefully chosen data which minimizes this error. This can be done most efficiently using codec specific error concealment tools, which treat encrypted data like any type of bitstream error. The JJ2000 version used in the experiments includes the patches and enhancements to JPEG2000 error concealment provided by [7].

2.2 Evaluation Methodology

We evaluated different settings for encoding as well as encryption. On the encoding side we looked at layer and resolution progression in JPEG2000. For encryption we have two basic question to answer: (A) Where is the most relevant information for the face recognition algorithms, and (B) what is the minimum amount of encryption required to protect the biometric face sample.

We will solely focus on the beginning of the JPEG 2000 codestream. The residual information towards the end of the codestream contains only fine texture data compared to the structural information at the beginning of the codestream. Further, face information depends on structural data, hence if encrypted structural data is given the residual data is practically worthless.

To detect the encryption location in the codestream with the biggest impact on the biometric recognition we are applying a "Sliding Window Encryption" [8–11]. In this method a window, i.e., a certain continuous percentage of the codestream size, is encrypted starting at a given offset. The sliding window uses a fixed window size for encryption and increases the offset from the beginning of the bitstream. For a proper evaluation of the

minimum encryption amount we will use an absolute encryption approach. Here we use a fixed offset (the beginning of the codestream) and will continually increase the encryption window size, this will be denoted 'increasing window encryption'.

The following setups will be used in the experiments: **small encryption window:** The size of the sliding window is fixed to 0.5% of the bitstream, the offset varies from 0% to 15% in steps of 1%; **large encryption window:** The sliding window is increased, but still fixed, in size to 4%, the offset is varied from 0% to 20% in 2% steps; **increasing encryption window:** The offset is fixed to the beginning of the codestream (0%) and the encryption window size increases from 1% to 15% in steps of 1%.

We will also assess the options available in the ordering of the JPEG 2000 bitstream (*resolution* and *layer progression*). Resolution progression starts with a downsampled version of the image and correction data for upsampling the resolution. The layer progression starts with the full resolution and the strongest frequencies, as in highest amplitude, in the wavelet domain, the ordering here is slightly more complex as coefficients from all wavelet sub-bands are mixed depending on sub-band (lowest first) and amplitude (largest first).

An illustration how this looks in practice is shown in Fig. 1 for layer progression mode and the small encryption window setup.

2.3 Face Recognition Methods

In face biometrics there are several techniques that can be used to fulfill a reliable verification or identification process, all of which been evaluated thoroughly in the literature, e.g. [12, 13]. Thus, we will describe the face recognition schemes selected for this study, including well-established traditional approaches as well as more recent ones based on deep convolutional neural networks, only briefly.

The considered methods based on traditional handcrafted local descriptors are using Local Binary Patterns (LBP) [14] and Multi-Block LBP (MBLBP) [15]. Both selected descriptors were extracted by using cells regions of size 14×14. The resulting feature vector, representing a face image, is represented by a histogram containing all single histograms of each cell region extracted before. The features are compared using the chi-square similarity measure.

The second class of applied methods contains three recent deep convolutional neural networks including ResNet-ArcFace (ArcFace) [16], MobileFaceNet (MobileFace) [17] and ShuffleFaceNet (ShuffleFace) [18]. In this case, the resulting features are compared using the cosine distance.

3 Experiments

There are a large number of face databases available that have been used in face recognition research. These databases vary in size, scope, purpose and thus include various illumination conditions (e.g. [19]), changes in facial expression or multiple poses (e.g. [20]). The Labeled Faces in the Wild (LFW) database [21], which we used in this study, is well known as a public benchmark for unconstrained face verification as the contained images have been collected from the web. The 13,233 face images are from 5,749 different identities, with large variations in pose, expression and illumination. As specified in [20] we have used a 10-fold split of 6000 face pairs in the experiments. All face images were aligned and cropped to the size of 112×112 pixel by using the RetinaFace detector [22] before subsequently applying JPEG2000 encryption followed by one of the described recognition schemes.

We will focus solely on the equal error rate (EER) while presenting the performance evaluation. Other considered measures, like the mean accuracy and the area under curve, agree with the EER and thus their presentation is skipped. A 50% EER is akin to guessing, meaning no information from the protected samples can be used in the biometric comparison. Furthermore, the EER, while not the most useful operation point on the receiver operator curve (ROC) from a practical standpoint, allows for easier and faster comparison to other results as shown in [23].

3.1 Evaluation

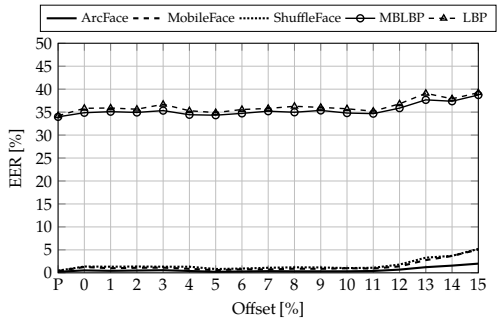
The results of the **small encryption window** experiments are shown in Fig. 2. The baseline without encryption, or, in cryptographic terms, a plain text experiment, is presented (labeled **P**). What is quite apparent from the figure is that the encryption window size of 0.5% has a negligible impact on the recognition performance and can not be considered secure.

The traditional methods, LBP and MBLBP, act similarly to the DL based methods. That is, an encryption that has an impact on the traditional methods has a similar effect on the DL based methods. This effect can also be seen for later experiments, see Figs. 3 and 4. The performance of the traditional methods is vastly inferior to the DL methods. But they are also much faster since they require no training. Given the matching behavior this means that the fast traditional methods can be used to gauge the impact of encryption. When using encryption to find the locations in the codestream, which are of primary interest for the performance of the biometric verification, this is a huge time saver.

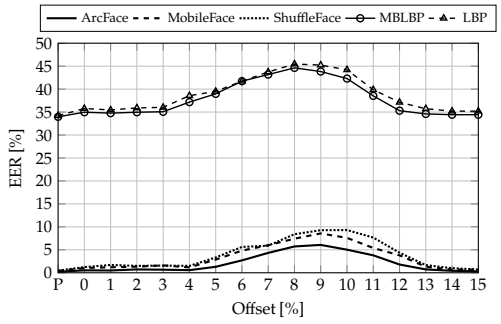
The impact of the small encryption window on the performance of the biometric system is based purely on the location where the encryption happens. And with the layer progression we can clearly see that this is a very specific location, roughly from offset 4% to 14%. This lies somewhere between the very coarse, basically global, structure and finer texture information. In Fig. 1 samples from this encryption method for layer progression mode can be seen. The impact of below and at 5% is mostly global noise, and from $\approx$ 12% onward we have a finer, texture level, noise.

Fig. 3 gives the result for the **large encryption window** experiment, showing EER over the offset where the encryption window starts. The larger window size is still insufficient to protect the data for the resolution progression. For layer progression we can see that this increased window size is sufficient to protect the important parts of the bitstream for all algorithms under test. The main difference compared to the small window encryption is that the encryption of the structural information, from 0% to 4%, also is sufficient for the protection of the biometric template. This would leave the coarse information which, according to the small window encryption experiment, is most

Figure 1: Sample from the faces in the wild database, Aaron Eckhart #1, with layer progression and a small encryption window at the given offset. The original unencrypted sample is also given (labled P).

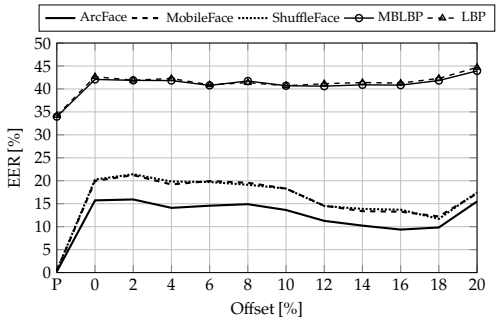


(a) Resolution progression with error correction.

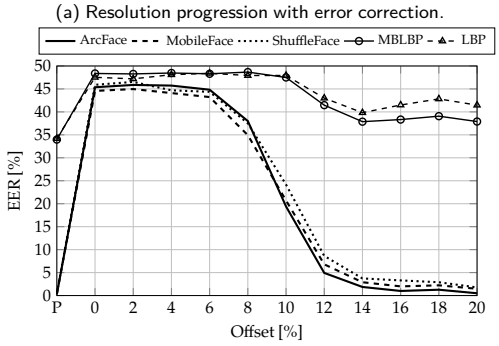


(b) Layer progression with error correction.

Figure 2: Equal error rates for the small encryption window with offset moving from 0 to 15% and the unencrypted baseline (**P**).



(a) Resolution progression with error correction.



(b) Layer progression with error correction.

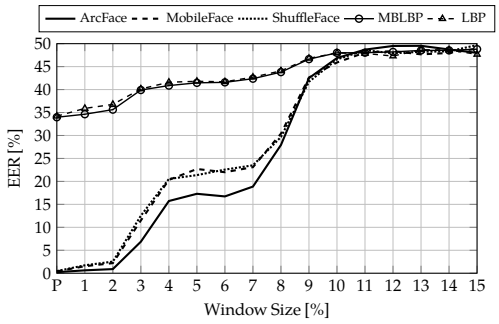
Figure 3: Equal error rates for the large encryption window with offset from 0 to 20% and the unencrypted baseline (**P**).

relevant for the biometric recognition in the clear. The reason for this is most likely that the removal of the basic structure makes the refinement information unusable to such an extent that the security is maintained. However, given that we know the information is in there, it is conceivable that it can be extracted to such an extent that an attack is made possible.

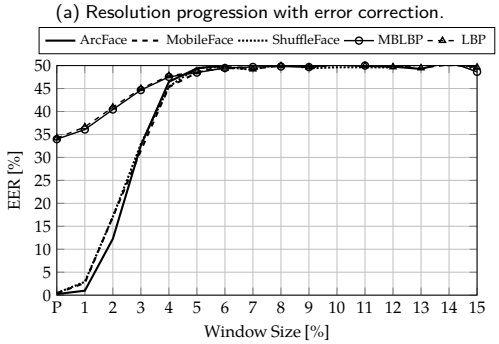
The results of the **increasing window encryption** are displayed in Fig. 4. Overall this experiment only confirms what was already found by the sliding window experiments. In the case of layer progression encoding the security is reached when encrypting the coarse structure that lies between 5% and 12%. For the resolution progression the same findings as before are also repeated. It is clear that the information grouped together in the layer progression mode is spread out among subbands of the wavelet decomposition used in JPEG2000. Due to this a much larger overall part of the code stream has to be encrypted for a similar security result compared to the layer progression.

4 Conclusion

The use of JPEG2000 encryption for the protection of face biometric samples was analysed, based on traditional and on deep



(a) Resolution progression with error correction.



(b) Layer progression with error correction.

Figure 4: Equal error rates for the increasing window encryption with a size of 1 to 15% and the unencrypted baseline (**P**).

learning based face recognition methods. The evaluation also took into account the different JPEG2000 encoding types, layer and resolution progression.

When storing facial biometric samples with JPEG2000 it is recommended to use the layer progression type. The relevant part for biometric face recognition is at around 4–12% of the total codestream. The most secure method for encryption is to start at the beginning and at least include this part of the codestream, i.e., encrypting the first 12%.

With respect to the encryption the traditional and deep learning based methods exhibit an identical behavior, the same information from an image was apparently used in the biometric comparison. Therefore, faster traditional methods can be used for analysis of selective encryption options. The effect on the DL based approaches will be identical but the time for an evaluation will be much reduced.

Acknowledgements

This project received funding from EU Horizon 2020 research program under grant agreement No. 690907 and from the Austrian Science Fund under project No. P27776.

References

[1] C. Rathgeb, A. Uhl, and P. Wild, *Iris Recognition: From Segmentation to Template Security*, ser. Advances in Information Security. 2013, vol. 59 (cit. on p. 2).

[2] *ISO/IEC 19794-5:2011 information technology — biometric data interchange formats — part 5: Face image data*, 2011 (cit. on p. 2).

[3] K. Delac, M. Grgic, and S. Grgic, "Effects of JPEG and JPEG2000 compression on face recognition," in *Proceedings of ICAPR 2005*, ser. LNCS, vol. 3687, 2005 (cit. on p. 2).

[4] A. K. John, O. O. Williams, and M. A. Adewale, "Evaluating the effect of jpeg and jpeg2000 on selected face recognition algorithms.," *International Journal of Modern Education & Computer Science*, vol. 6, no. 1, 2014 (cit. on p. 2).

[5] D. Engel, T. Stütz, and A. Uhl, "A survey on JPEG2000 encryption," *Multimedia Systems*, vol. 15, no. 4, 2009. DOI: http://dx.doi.org/10.1007/s00530-008-0150-0 (cit. on p. 2).

[6] F. Dufaux, S. Wee, J. Apostolopoulos, and T. Ebrahimi, "JPSEC for secure imaging in JPEG2000," in *Applications of Digital Image Processing XXVII*, vol. 5558, 2004 (cit. on p. 2).

[7] T. Stütz and A. Uhl, "On JPEG2000 error concealment attacks," in *Advantages in Image and Video Technology: Proceedings of the 3rd Pacific-Rim Symposium on Image and Video Technology, PSIVT '09*, ser. Lecture Notes in Computer Science, (cit. on p. 2).

[8] M. Rieger, J. Hämmerle-Uhl, and A. Uhl, "Efficient iris sample data protection using selective jpeg2000 encryption of normalised texture," in *Proceedings of the 6th International Workshop on Biometrics and Forensics (IWBF'18)*, 2018 (cit. on p. 2).

[9] ——, "Selective JPEG2000 Encryption of Iris Data: Protecting Sample Data vs. Normalised Texture," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19)*, 2019. DOI: 10.1109/ICASSP.2019.8683196 (cit. on p. 2).

[10] M. Draschl, J. Hämmerle-Uhl, and A. Uhl, "Sensor dependency in efficient fingerprint image protection using selective jpeg2000 encryption," in *Proceedings of the 5th International Workshop on Biometrics and Forensics (IWBF'17)*, 2017 (cit. on p. 2).

[11] S. Shekhawat, H. Hofbauer, B. Prommegger, and A. Uhl, "Efficient fingervein sample image encryption," in *Proceedings of the 8th International Workshop on Biometrics and Forensics (IWBF'20)*, Shortlisted for best paper award, 2020 (cit. on p. 2).

[12] Y. Xu, Z. Li, J. Yang, and D. Zhang, "A survey of dictionary learning algorithms for face recognition," *IEEE access*, vol. 5, 2017 (cit. on p. 3).

[13] R. Jafri and H. R. Arabnia, "A survey of face recognition techniques," *journal of information processing systems*, vol. 5, no. 2, 2009 (cit. on p. 3).

[14] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 12, 2006 (cit. on p. 3).

[15] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, "Face detection based on multi-block lbp representation," in *International conference on biometrics*, Springer, 2007 (cit. on p. 3).

[16] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019 (cit. on p. 3).

[17] S. Chen, Y. Liu, X. Gao, and Z. Han, "Mobilefacenets: Efficient cnns for accurate real-time face verification on mobile devices," in *Biometric Recognition*, 2018 (cit. on p. 3).

[18] Y. Martindez-Diaz, L. S. Luevano, H. Mendez-Vazquez, M. Nicolas-Diaz, L. Chang, and M. Gonzalez-Mendoza, "Shufflefacenet: A lightweight face architecture for efficient and highly-accurate face recognition," in *The IEEE International Conference on Computer Vision (ICCV) Workshops*, 2019 (cit. on p. 3).

[19] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, no. 6, 2001 (cit. on p. 3).

[20] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The cas-peal large-scale chinese face database and baseline evaluations," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 38, no. 1, 2007 (cit. on p. 3).

[21] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, 2007 (cit. on p. 3).

[22] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-stage dense face localisation in the wild," *CoRR*, vol. abs/1905.00641, 2019. arXiv: 1905.00641 (cit. on p. 3).

[23] H. Hofbauer and A. Uhl, "Calculating a boundary for the significance from the equal-error rate," in *Proceedings of the 9th IAPR/IEEE International Conference on Biometrics (ICB'16)*, 2016 (cit. on p. 3).