

© Springer Verlag. The copyright for this contribution is held by Springer Verlag. The original publication is available at www.springerlink.com.

Exploring Compression Impact on Face Detection using Haar-like Features

Peter Elmer, Artur Lupp, Stefan Sprenger, René Thaler and Andreas Uhl¹

University of Salzburg, Department of Computer Sciences, Jakob Haringerstr. 2, 5020
Salzburg, Austria,
`andreas.uhl@sbg.ac.at`

Abstract. The main goal in our experimental study was to explore the impact of image compression on face detection using Haar-like features. In our setup we used the JPEG, JPEG2000 and JPEG XR compression standards to compress images from selected databases at given compression ratios. We performed the face detection using OpenCV on the reference images from the database as well as on the compressed images. After the detection process we compared the detected areas between the reference and the compressed image gaining the average coverage, false positive and false negative areas. Experimental results comparing JPEG, JPEG2000 and JPEG XR are showing that the average coverage of the detected face area differ between 79,58% in the worst and 99,61% in the best case. The false negative (not covered) areas range between 0,33% and 19,75% and false positive (fallout) areas between 0,38% and 9,45%. We conclude that the JPEG compression standard is performing worse than JPEG2000 and JPEG XR while both latter providing quite equal and good results.

1 INTRODUCTION

While the research in the field of face recognition is blooming, face detection receives much less attention. Nevertheless face detection, widely used for interactive user interfaces or as a feature in cameras, is the very first and necessary stage for many automated and semi-automated face recognition systems. Given that, the tasks of detecting faces for recognition are becoming required more frequently - e.g. in security systems at airports for preventing acts of terrorism. In order to detect a face it is necessary to differentiate between the background and the desired area where the face is. The huge quantities of visual data collected and stored makes the application of lossy image compression algorithms ubiquitous. The impact of these algorithms on pattern recognition tasks like face detection is often neither well investigated nor well understood.

Image compression algorithms are classically either optimised with respect to human perception (e.g. the JPEG default quantisation (Q-)table) or with respect to rate-distortion criteria (e.g. Tier-2 coding in JPEG 2000 or design of a specific Q-table for JPEG). For applications in pattern recognition, optimisation with respect to these criteria is not necessarily the optimal solution. For example,

in [17] the JPEG Q-table is tuned for application in the pattern recognition context by emphasising middle and high frequencies and discarding low frequencies (the standard JPEG Q-table is rotated by 180 degrees). JPEG Q-table optimisation has already been considered in face recognition [16] which leads to superior recognition performance as compared to the standard matrix. A further example is the optimisation of JPEG 2000 Part 2 wavelet packet decomposition structures with respect to optimising iris recognition accuracy [26] which provides better results compared to rate-distortion optimised wavelet packet structures. These observations raise the question and severe doubt if compression algorithms exhibiting better rate-distortion performance are indeed better in a specific pattern recognition context.

With respect to standardization of image compression in biometrics, the ISO/IEC IS 19794 represents the most relevant standard, recommending JPEG 2000 exclusively for lossy compression. Also, 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) supports only JPEG 2000 for applications tolerating lossy compression. Apart from standardization, a variety of independent studies dealing with compression and the respective impact on biometric recognition performance exist: E.g., iris compression [11], fingerprint compression [12, 13] and 3D face recognition [14]. Also for face image compression and its impact on face recognition performance a significant corpus of research exists [15], considering e.g. the impact of JPEG [16], JPEG 2000 [18, 25], SPIHT [19], and H.264 [20]. However, systematic investigations on the effect of lossy compression on **face detection** robustness and accuracy have remained elusive. The only work in this direction is restricted to JPEG [21] and also considers tracking and recognition. Most work on face detection and compression deals with face detection techniques in the compressed domain, i.e. detecting faces analysing the bitstream only without the necessity for full decoding (e.g. from JPEG and MPEG [22, 23], H.264 [24], and wavelet-based data).

In the seminal work of Viola and Jones [5, 6] they describe a way for rapid object detection using simple distinctive features. Haar-like features have scalar values representing the differences between two rectangular regions in average intensities. Using these features faces are distinguished from the background and the desired area where the face gets detected. Due to the simple set of classifiers used the general error rate is quite high. For better detection rates, an extension of these Haar-like features, i.e. a specially trained cascade set introduced by Lienhart and Maydt [8] can be used. The OpenCV [9] Library offers this feature set for face object detection.

In this work, we use OpenCV Viola-Jones face detection with the Haar Cascade File created by Lienhart and three widely used ISO still image coding standards: JPEG, JPEG2000 and JPEG XR. The main aim is to investigate the robustness of face detection to the three types of compression and to answer the question if face detection robustness matches to measured image quality in terms of PSNR. In our study three different sets of images were converted using 10 different compression ratios with each compression standard. After the

compression, face detection was applied on the reference images from the image sets and the compressed images. Then we compare the detected areas between the reference images from the data sets and the compressed ones showing the average, false negative and false positive coverage. Experimental results achieved with this method show that JPEG is performing poor with low bitrate pictures and high compression ratios whereas JPEG2000 and JPEG XR show consistent performance. In the following sections, we first describe the compression standards we used. Thereafter the data sets, employed software and setup used in our experiments are described followed by results and a conclusion.

2 COMPRESSION STANDARDS

2.1 JPEG

JPEG (Joint Photographic Experts Group ISO/IEC IS 10918-1) [10], despite its age, is still one of the most used standards in digital imaging technology. The JPEG compression divides the image data into blocks of 8x8 pixels. To each of these blocks a discrete cosine transform (DCT) is applied. Depending on the quality factor, quantization is applied on the coefficients of the DCT. Then the data is compressed using Huffman encoding. Due to the quantization reconstruction errors occur and are distributed over an entire 8 x 8 pixels block, leading to blocking artifacts.



(a) Face detection applied on JPEG images. The left image is the reference image. The image on the right is compressed with a compression ratio of 2. As mentioned, the area where the face is detected differs a little bit from the reference image. [Image is Part of LFW Database]



(b) This image was compressed with ratio of 100. Due its high compression leading to blocking artifacts, the detection of a face was not possible. [Image is Part of LFW Database]

Fig. 1: JPEG Compression

These artifacts have an impact on the face detection rate. See [Figure 1a] for an example of face detection applied on a JPEG image. The lower the quality

factor the higher the compression leading to more blocking artifacts which can abstract an image with faces to a point where the face detection is inaccurate or simply not possible at all. See [Figure 1b] for an example of face detection applied on a JPEG image.

2.2 JPEG2000

JPEG2000 was created as well by the JPEG committee (ISO/IEC IS 15444-1) with the intention of superseding the DCT based approach used in the JPEG compression standard with a wavelet based method. The JPEG2000 compression standard is highly flexible offering a superior compression performance than JPEG, a multiple resolution representation and the possibility to either use a lossless or lossy compression. During compression in JPEG2000 the image is, like in JPEG, optionally partitioned into rectangular non overlapping blocks (tiles). The tiles can be compressed independently as they were entirely distinct images. Wavelet transform, quantization and entropy coding are applied independently in these blocks. Splitting the image into tiles reduces memory requirements and since they are reconstructed independently it is possible to decode specific parts of the image instead of the whole image. This however is required only for large images and induces artifacts.



(a) As in [Figure 1a], these pictures show the reference image (left) and the same image compressed with the compression ratio of 2 (right). There are just minor differences in the detected areas.

(b) Image with a compression ratio of 100. If we Compare [Figure 2b] with [Figure 1b], there are significant less blocking artifacts using JPEG2000.

Fig. 2: Face detection applied on JPEG2000 compressed image [Image is Part of LFW Database]

Coding of coefficient subsets (“code-blocks”) results in a code stream after being coded by the arithmetic coding stage which is called Tier-1. Tier-1 coding produces a code steam containing first the data with the greatest distortion reductions achieved through a fractional bit-plane coder. The last stage of coding is the reorganization of the code stream. This process codes the auxiliary data which is needed to identify the content of the quality layers into the code stream.

The auxiliary data is stored in the packet header of the code blocks. This header contains the information for the code block whether or not the block contributes to the quality layer, the number of the encoding passes, length of the encoded data, and the number of the magnitude bit planes. This allows features like quality and resolution progressiveness. It is to note that the superior compression as well as the high flexibility of this standard have to be paid with restrictions in terms of computational performance since run-time can be increased by a factor of 10 as compared to JPEG. As visible in [Figure 2b] there are (of course) much less blocking artifacts compared to JPEG compressed images [Figure 1b], while pictures compressed with low compression ratios are quite similar to the original [Figure 2a].

2.3 JPEG XR

JPEG XR is the most recent still image coding standard from the JPEG committee (ISO/IEC IS 29199-2), published in early 2009, we used to evaluate the compression impact on. Primarily targeting extended range imagery initially, JPEG XR finally managed in getting high image quality almost equivalent to JPEG2000. The advantages compared to JPEG2000 are on one hand primary the lower complexity allowing us to archive almost identical quality during compression while needing lower computational resources as well as storage capacity, on the other hand also providing high dynamic range support. As JPEG2000, JPEG XR is offering better lossy compression ratios in comparison to JPEG for encoding an image with identical quality, the possibility to compress lossless, support for more color accuracy and transparency map support and meta data support. It is one of the most recent still image coding standards available. Due to its lower complexity compared to JPEG2000 we wanted to evaluate if there is also a difference in terms of face detection. In the work of Horvat, Stögner and Uhl, JPEG XR was analyzed on Iris Recognition Systems [7] and they concluded that JPEG XR performed better than JPEG2000.



Fig. 3: Face detection applied on images compressed JPEG XR. From left to right the compression ratios are: 1 (reference image), 2 and 100. Image quality is almost identical to JPEG2000. [Image is Part of LFW DB]

3 EXPERIMENTAL SETUP AND RESULTS

This section describes our datasets, the software and setups we used. At the end of this section we will provide the results of our experiments.

3.1 Used Databases

Labeled Faces in the Wild [1] The Labeled Faces in the Wild Database set contains more than 13000 images of faces collected from the web. We randomly selected 431 of these .jpg images with the average quality factor of 75. Most of these images show motives of people during events or speeches. The resolution of the images differ, starting from 233 x 409 and going up to 410 x 450 pixels. This data set of images contains motives with more than one face. Due to that fact there is a high possibility that small faces will not be detected at high compression ratios. See [Figure 2b] for an example.

BIOID Face Database [2] This dataset consists of 1521 grey level images with a resolution of 284 x 386 pixels recorded under natural conditions so that the test set is featuring a large variety of illumination, background as well as face sizes. The images are labeled "BioID_XXXX.pgm" where the characters XXXX are replaced by the index of the current image (with leading zeros).

Bao Face Database [3] The Bao Face Database consists of 221 .jpg images with mostly people from Asia. The resolution differs, starting from 250 x 205 and going up to 1836 x 1190 pixels. Some images were altered or cropped. The images are labeled "x.jpg" where x is a number between 1 and 221. As this set of images contains a lot of group picture motives it is very likely that with increasing compression ratios face detection will miss faces. Therefore the false negative areas will tend upwards as compression ratio is increasing.

Please note that with these datasets, we are able to assess eventual influence of double compression. Since the BIOID dataset comes in uncompressed grey level format, similar behaviour to one of the other two datasets confirms that no double compression effects are observed. In the plots we will refer to the datasets with abbreviations: "LFW DB" for Labeled Faces in the Wild, "BIOID DB" for BIOID Face Database and "BaoDB" for the Bao Face Database

3.2 Setup

The images from the data sets were compressed with JPEG, JPEG2000 and JPEG XR. We implemented the standard libraries and converted the images with given compression ratios. The Software we used in our setup to convert the image data was ImageMagick with default settings. To obtain a specific compression ratio, a divide-and-conquer algorithm was used to determine the quality parameter for the specific compression ratios. We took a reference image

out of the dataset and obtained its size. Then we applied the compression on this file and checked if the compressed file size is the target file size we wanted to have. The target file size depends on the compression ratio. With a compression ratio of 50 the target file size should be 1/50 of the original file size. To obtain the desired file size for our images, we changed the quality parameters for the compression. Our compression ratios were 1, 2, 5, 10, 12.5, 16.67, 25, 33.33, 50, 100 for this experiment. In the beginning the ratios are very small, that is to evaluate eventual double compression effects.

After the image compression, we performed face detection using the OpenCV library with the Haar Cascade File (haarcascade_frontalface_alt.xml [4]) created by Lienhart. During the face detection the reference image was used first. We gained the desired areas as y and x coordinates where x ranges from 0 to the maximal width and y ranges from 0 to the maximal height of the image. Each detected face has four parameters. These parameters are the four corners of the detected face area. We used these parameters to compare them with the ones achieved during the face detection of the compressed images. While comparing the areas of the reference image with the areas of the compressed images we gained three values. The first one is the average coverage of the areas. The more of the reference area where the face was detected in the original was covered in the compressed image, the higher this value gets (denoted as “**cov**” in the plots). The second value describes the false negative areas that were covered in the reference but not on the compressed image (denoted as “**notc**” in the plots). The third value describes the false positive areas that were covered in the compressed image but not in the reference image (denoted as “**morc**” in the plots). These values were saved for each picture in a database. With this data we could compute the average coverage, average false positive and average false negative for our images. For each dataset we gained three different plots [See Figures 5 to 7 for Bao Face Database, 8 to 10 for BIODID Database and 11 to 13 for Labeled Faces in the Wild] for each compression standard used. Each of these plots [Figures 5 to 13] is showing the average coverage, average false positive and average false negative for the specific compression ratio.

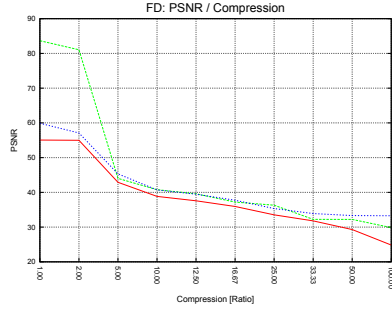
3.3 Results

First, we display the average rate-distortion performance of the three compression standards on the considered datasets. In [Figure 4a and 4b] we observe the classical and expected behaviour: Especially for high compression ratios, JPEG performance gets very poor, while JPEG2000 and JPEG XR are close with slight advantages for JPEG2000. For enabling a fair comparison, the same Haar Cascade file [4] was used during all face detection experiments. The following plots show the coverage using the y-axis value on the left side where a higher value is better. The false negative areas (notc) and the false positive areas (morc) are using the right y-axis where a lower value is better. The x-axis is showing the used compression ratio. Looking at [Figures 5 to 7], JPEG is performing very poor because the images in the Bao Face Database are patchy and already pre compressed in JPEG. Comparing the results from the images compressed using

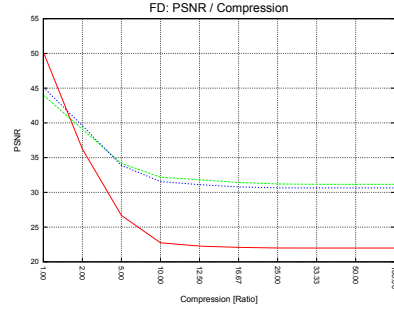
JPEG2000 [Figure 6] and JPEG XR [Figure 7] we can see that the results are quite alike but JPEG XR is performing better if we look at the numbers. At compression ratio 100, images compressed using JPEG2000 [Figure 6] have an average coverage of 93.50, false negative 5.86 and false positive 7.15. Detecting faces on images compressed with JPEG XR [Figure 7] is slightly better. The average coverage is 94.06, false negative 5.29 and false positive 5.16. This set of images shows the highest false positive and false negative values we achieved in our experiments. This is most likely due the group picture motives in the images. With higher compression and emerging blocking artifacts faces especially small ones will not be detected. [Figures 8 to 10] show us the results for the BIOID Database. As all the images provided by the data set have the same aspect ratio and only one face to detect, the over all performance is good for all compression standards used. It also seems, that there is no difference if the detected images are color or gray scale images. Looking at [Figure 8] which shows us the results of the face detection using JPEG, we can see a slightly higher area with false positives at compression ratio 100. If looking at [Figure 9] using JPEG2000 and [Figure 10] using JPEG XR compression, JPEG XR performs better. The BIOID data set achieves the best results with a good over all coverage, low false positives and false negatives. That is because each of the images in this data set shows one frontal face only. The Labeled faces in the Wild Database (JPEG pre-compressed as well), shows a quite similar trend as the rest of the datasets, see [Figures 11 to 13], JPEG2000 [Figure 12] and JPEG XR [Figure 13] achieving better results compared to JPEG [Figure 11]. The results of the images compressed using JPEG XR [Figure 13] are slightly better than JPEG2000 [Figure 12]. Despite the differences between the three data sets there are similarities when analyzing the results. The Bao Face Database as well as the Faces in the Wild Database have higher false negatives and false positives compared to the BIOID Database due images with more faces on them.

4 CONCLUSION

Our experiment demonstrates that in the context of face detection using sets of Haar-like features, JPEG XR and JPEG2000 are superior to JPEG, especially when the images are compressed at high ratios. Due to the significant lower computational demand and slightly better results in face detection JPEG XR is a good alternative to JPEG2000 in all considered settings. In order to focus on the differences between JPEG XR and JPEG2000 we show \pm one standard deviation in addition to the mean values [Figures 14 to 16] which confirm the former observation. Concerning the relation between PSNR and detection accuracy we have found that JPEG is correctly predicted to be the worst algorithm by its PSNR values, however, the slight advantages of JPEG XR over JPEG 2000 are not correctly predicted by PSNR, in contrary [Figure 4]. It has also turned out that a slight JPEG precompression does not influence the performance and the ranking of the three compression algorithms when considering the similarity of the results obtained on pre-compressed and not pre-compressed datasets.



(a) BIODID Database



(b) Bao Database

Fig. 4: PSNR / Compression Ratio

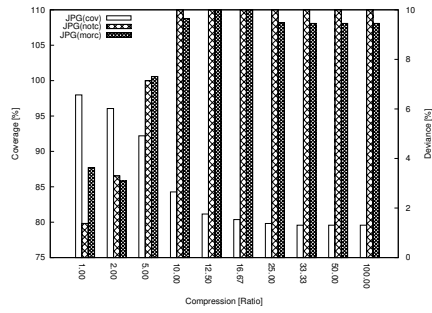


Fig. 5: BaoDB - JPEG. At compression ratio 100, avg cov: 79.58, FN: 19.75, FP: 9.45

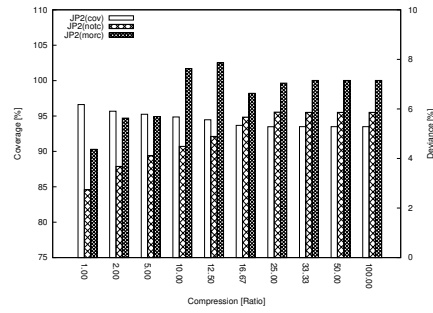


Fig. 6: BaoDB - JPEG2000. At compression ratio 100, avg cov: 93.50, FN: 5.86, FP: 7.15

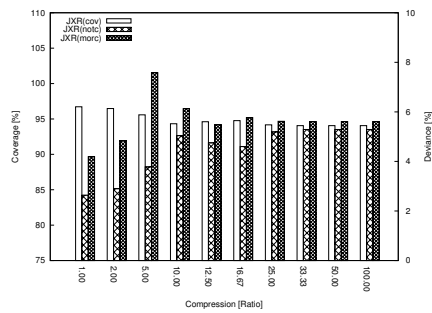


Fig. 7: BaoDB - JPEG XR. At compression ratio 100, avg cov: 94.06 FN: 5.29, FP: 5.61

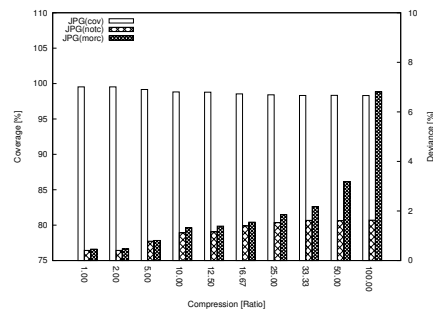


Fig. 8: BIODID DB - JPEG. At compression ratio 100, avg cov: 98.31, FN: 1.63, FP: 6.81

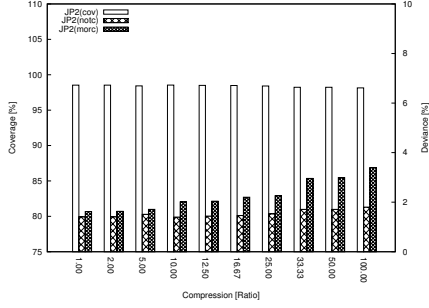


Fig. 9: BIOID DB - JPEG2000. At compression ratio 100, avg cov: 98.14, FN: 1.80, FP: 3.42

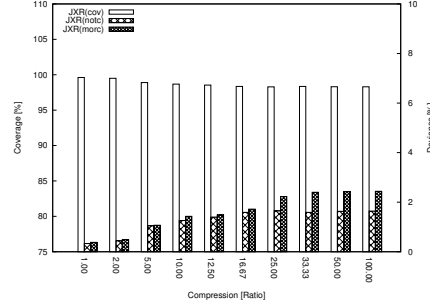


Fig. 10: BIOID DB - JPEG XR. At compression ratio 100, avg cov: 98.30, FN: 1.64, FP: 2.44

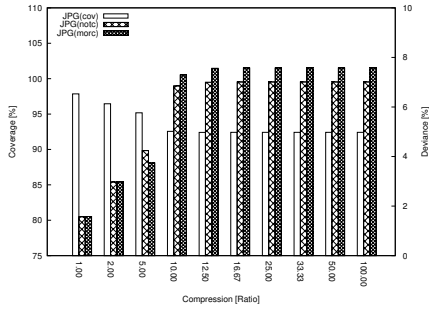


Fig. 11: LFW DB - JPEG. At compression ratio 100, avg cov: 92.42, FN: 7.01, FP: 7.57

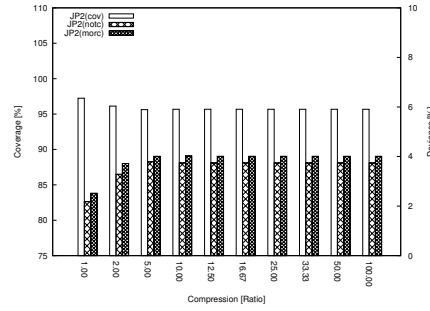


Fig. 12: LFW DB - JPEG2000. At compression ratio 100, avg cov: 95.68, FN: 3.74, FP: 4.01

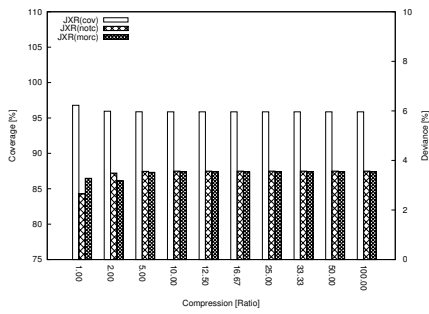


Fig. 13: LFW DB - JPEG XR. At compression ratio 100, avg cov: 95.87, FN: 3.57, FP: 3.54

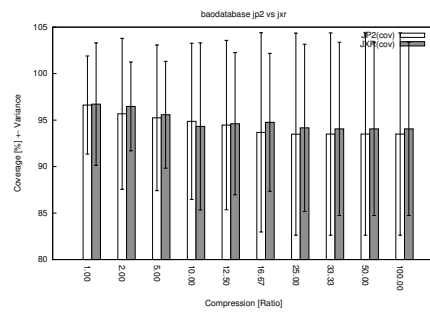


Fig. 14: Bao DB - JPEG XR & JPEG2000, Coverage with standard deviations

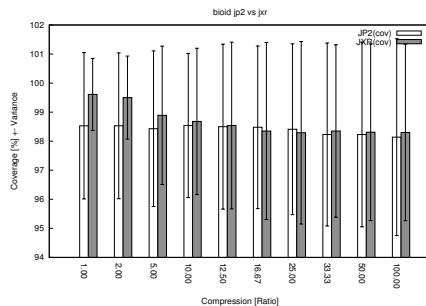


Fig.15: BIOD DB - JPEG XR & JPEG2000, Coverage with standard deviations

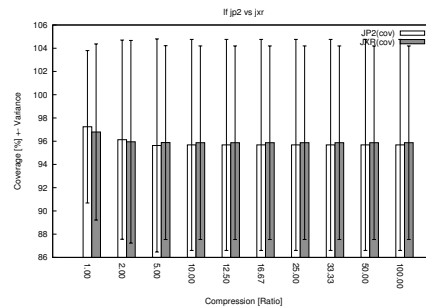


Fig.16: LFW DB - JPEG XR & JPEG2000, Coverage with standard deviations

References

1. Gary B. Huang, Manu Ramesh, Tamara Berg and Erik Learned-Miller. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. University of Massachusetts, Amherst, October 2007, Pages 07-49
2. O. Jesorsky, K. Kirchberg, R. Frischholz, Face Detection Using the Hausdorff Distance, In J. Bigun and F. Smeraldi, editors, Audio and Video based Person Authentication - AVBPA 2001, pages 90-95. Springer, 2001
3. R. Frischholz. Bao face database at the face detection homepage, <http://www.facedetection.com>, Last accessed 9/26/2014
4. Stump-based 20x20 gentle adaboost frontal face detector, Created by Rainer Lienhart, https://github.com/Itseez/opencv/blob/master/data/haarcascades/haarcascade_frontalface_alt.xml. Last accessed 9/23/2014
5. Viola, P.; Jones, M., "Rapid object detection using a boosted cascade of simple features," Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on , vol.1, no., pp.I-511,I-518 vol.1, 2001
6. Viola, Paul and Jones, Michael J.: Robust Real-Time Face Detection, International Journal of Computer Vision, Volume 57, Issue 2, 137154 (2004)
7. Kurt Horvath, Herbert Stögner, and Andreas Uhl. 2011. Effects of JPEG XR compression settings on iris recognition systems. In Proceedings of the 14th international conference on Computer analysis of images and patterns - Volume Part II (CAIP'11), Vol. Part II. Springer-Verlag, Berlin, Heidelberg, 73-80.
8. Lienhart, R.; Maydt, J., "An extended set of Haar-like features for rapid object detection," Image Processing. 2002. Proceedings. 2002 International Conference on , vol.1, no., pp.I-900,I-903 vol.1, 2002
9. Bradski, G., "The OpenCV Library", Dr. Dobb's Journal of Software Tools, 2000. Available from: <http://drdobbs.com/opensource/184404319>
10. Wallace, G.K., "The JPEG still picture compression standard," Consumer Electronics, IEEE Transactions on , vol.38, no.1, pp.xviii,xxxiv, Feb 1992
11. S. Rakshit and D.M. Monro, An Evaluation of Image Sampling and Compression for Human Iris Recognition, IEEE Transactions on Information Forensics and Security, 2007, vol. 2, no. 3, p. 605-612

12. M.A. Figueroa-Villanueva and N.K. Ratha and R.M. Bolle, A comparative performance analysis of JPEG2000 vs. WSQ for fingerprint compression, Proceedings of AVBPA, LNCS, vol. 2688, Springer Verlag, p. 385-392, 2003
13. R.C. Kidd, Comparison of wavelet scalar quantization and JPEG for fingerprint image compression, Journal of Electronic Imaging, 1995, vol. 4, no. 1, p. 31-39
14. L. Granai, J.R. Tena, M. Hamouz and J. Kittler, Influence of compression on 3D face recognition, Pattern Recognition Letters, vol. 30, no. 8, p. 745-750
15. K. Delac, S. Grgic and M. Grgic, Image Compression in Face Recognition - a Literature Survey, Recent Advances in Face Recognition, I-Tech, p. 236-250, 2008
16. G.-M. Jeong, C. Kim, H.-S. Ahn and B.-J. Ahn, JPEG Quantization Table Design for Face Images and Its Application to Face Recognition, 2006, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Science, p. 2990 - 2993, vol. E69-A, no. 11
17. M. Chen, S. Zhang, M.A. Karim, Modification of standard image compression methods for correlation-based pattern recognition, 2004, Optical Engineering, p. 1723-1730, vol. 43, no. 8
18. K. Delac, S. Grgic and M. Grgic, Face recognition in JPEG and JPEG2000 compressed domain, Image and Vision Computing, vol. 27, p. 1108-1120, 2009
19. M. Kamasack and B. Sankur, Face recognition under lossy compression, Proceedings of the International Conference on Pattern Recognition and Information Processing (PRIP'99), 1999, p. 27-32
20. B. Klare and M. Burge, Assessment of H.264 Video Compression on Automated Face Recognition Performance in Surveillance and Mobile Video Scenarios, Biometric Technology for Human Identification VII, Proceedings of SPIE, vol. 7667, 2010, p. 76670X
21. P. Korshunov and W.T. Ooi, Video quality for face detection, recognition, and tracking, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 7, no. 3, article 14, 2011
22. H. Luo and A. Eleftheriadis, On face detection in the compressed domain, Proceedings of the ACM International Conference on Multimedia 2000, 2000, p. 285-294
23. P. Fonseca and J. Nesvadha, Face detection in the compressed domain, Proceedings of the IEEE International Conference on Image Processing (ICIP'04), 2004, p. 2015 - 2018
24. S.-S. Zhuang and S.-H. Lai, Face detection directly from h.264 compressed video with convolutional neural network, Proceedings of the IEEE International Conference on Image Processing (ICIP'09), 2009, p. 2485 - 2488
25. G.W. Quinn and P. J. Grother, Performance of Face Recognition Algorithms on Compressed Images, NIST Interagency Report 7830, Information Technology Laboratory, The National Institute of Standards and Technology, 2011
26. J. Hämmerle-Uhl, M. Karnutsch, A. Uhl, "Evolutionary Optimisation of JPEG2000 Part 2 Wavelet Packet Structures for Polar Iris Image Compression", In Proceedings of the 18th Iberoamerican Congress on Pattern Recognition (CIARP'13), pp. 391-398, Havana, Cuba, Springer LNCS, 8258, November 20-23, 2013