

Towards Face Recognition in JPEG2000 Compressed Domain

Kresimir Delac, Mislav Grgic, Sonja Grgic
University of Zagreb, Faculty of Electrical Engineering and Computing
Department of Wireless Communications
Unska 3 / XII, HR-10000 Zagreb, Croatia
Phone: + 385 1 6129 851, Fax: + 385 1 6129 568, E-mail: kdelac@ieee.org

Keywords: face recognition, compressed domain, JPEG2000, PCA, LDA, ICA

Abstract - The goal of this research is to examine the feasibility of implementing face recognition algorithms directly into JPEG2000 compressed domain. By exploring some standard face recognition algorithms' behavior when fed with standard JPEG2000 DWT coefficients (CDF 9/7 wavelet), we will prove that recognition can be done directly in JPEG2000 compressed domain, avoiding inverse discrete wavelet transform (IDWT). Such an approach would consequently enable the use of compressed images in recognition purposes, thus reducing both computational time and storage requirements. Our experimental results indicate that face recognition performance in JPEG2000 compressed domain is comparable, or even better in some cases, than face recognition performance in pixel domain.

1. INTRODUCTION

Face recognition has gained much attention in recent years and has become one of the most successful applications of image analysis and understanding. A general statement of the problem can be formulated as follows [1]: Given still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces. Images are expected to be increasingly stored in compressed form. Although JPEG is currently a *de facto* standard for storing still images, JPEG2000 [2] seems to give promising results and will probably replace JPEG in near future. No previously published work in face recognition has looked into JPEG2000 compressed domain face recognition, although there was some work done and some promising results reported in JPEG (i.e. DCT) compressed domain [3], [4], [5]. Wavelets were previously used in face recognition [6], [7], [8], but never Cohen-Daubechies-Feauveau (CDF) 9/7 wavelet, which is a part of JPEG2000 standard.

In this paper we explore the performance of some well known face recognition algorithms in JPEG2000 compressed domain. Compressed domain means that instead of decompressing the compressed images and then using (distorted) pixel values as input to face recognition methods, transformed coefficients are used as inputs. The decoding process should be interrupted after the entropy decoding and the obtained DWT coefficients used as inputs to classification methods. This way it is possible to achieve large computational time saves by avoiding the inverse DWT, and storage requirements saves by being able to store the images in a compressed form. We will give a detailed analysis of the influence that CDF 9/7 wavelet decomposition and coefficient coding has on performance results for known face recognition

algorithms, namely, Principal Component Analysis (PCA) [9], Independent Component Analysis - Architecture 2 (ICA) [10] and Linear Discriminant Analysis (LDA) [11].

To ensure that our results are easily comparable and reproducible, we will use FERET face database [12] along with its standard test sets and accompanying protocol.

We will show that compressing face images with JPEG2000 and using DWT coefficients in recognition does not deteriorate performance. This continues the study reported in [13] where we showed that compression (JPEG and JPEG2000) does not deteriorate face recognition performance. In [13] we used the compressed images and the distorted pixel values were the input to face recognition methods. We actually showed that artifacts introduced by coefficient quantization and entropy coding (e.g. blocking artifacts in JPEG and smearing in JPEG2000) do not deteriorate recognition performance. In fact, in some cases compression even improves performance. This effect was most often observed for bitrates between 0.2 and 0.8 bpp (bits per pixel). Wijaya et al. in [14] performed face verification on images compressed to 0.5 bpp by JPEG2000 and showed that high recognition rates can be achieved using correlation filters. Both these papers give reason to believe that implementing face recognition directly into the compressed domain is justifiable.

Our experimental results in this paper indicate that face recognition in JPEG2000 compressed domain would be comparable or even better than face recognition in pixel domain. It appears that JPEG2000 compression can effectively remove non-essential information from images, thus increasing face recognition performance results in some cases. We will also show a surprising performance increase in some algorithms for the different illumination problem. These results additionally justify possible implementation of face recognition algorithms in JPEG2000 compressed domain, before the inverse DWT, thus reducing computation time.

Another interesting thing about using wavelets in face recognition in general is their possible connection to human visual system. As stated in [7], according to psychovisual research there is much evidence that human visual system processes images in a multiscale way. Research evidence in neurophysiology and psychology is consistent with the notion that the visual system analyses input at several spatial resolution scales [15]. Along these lines of thought, multiresolution spatial frequency preprocessing (such as DWT) of faces is justified. An image is represented as a weighted combination of basis

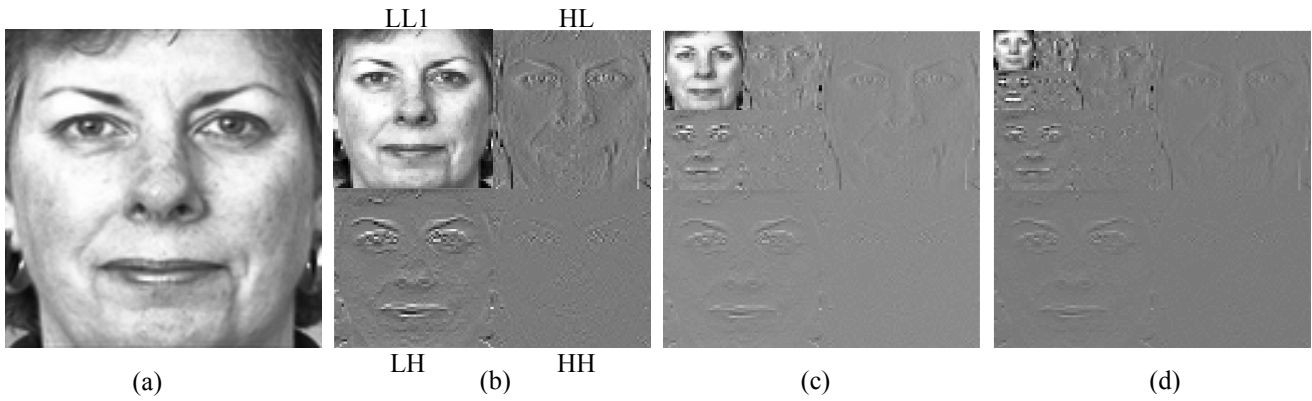


Fig. 1. DWT examples: (a) original image; Decomposed images: (b) one level decomposition, (c) two level decomposition, (d) three level decomposition.

functions (wavelets), where high frequencies carry detail information and low frequencies carry coarsely approximated information.

The rest of this paper is organized as follows: section 2 gives a brief introduction to wavelets (with connection to JPEG2000 standard) and their use in face recognition so far, section 3 describes the experimental setup, in section 4 we discuss and analyze our results and section 5 concludes the paper and gives some hint of our further work.

2. WAVELETS

2.1 Wavelets and image compression

Wavelet transform (WT) has been a very popular and commonly used tool for image analysis over the last decade or so and, as such, has become the part of JPEG2000 standard. JPEG2000 uses Cohen-Daubechies-Feauveau (CDF) 9/7 wavelet as a default wavelet for lossy compression. It is a biorthogonal wavelet that uses filters with dissimilar filter lengths for decomposition and reconstruction. Detailed analysis of CDF 9/7 wavelet and JPEG2000 standard is out of scope of this paper and an interested reader is referred to [2], [16], [17], [18] for details on these issues. Wavelets in general are mathematical basis functions that are used to represent other functions. WT represents an image as a combination of wavelet functions (wavelets) with different locations and scales [16].

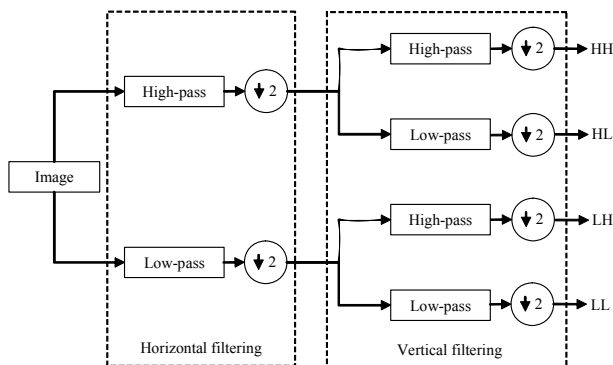


Fig. 2. Illustration of 2-D wavelet transform.

The implementation of two-dimensional WT is carried out by applying a one-dimensional transform to the rows of the original image data and the columns of the row-transformed data respectively. We can say that WT performs multiresolution image analysis (a simultaneous image representation on different resolution and quality levels).

It is important to mention that a wavelet transform is in practice created by passing the image through a series of filter bank stages. An illustration of one stage (one decomposition) is shown in Fig. 2 and an image example of this in Fig. 1(b). An image is first filtered in the horizontal direction. The high-pass filter and the low-pass filter are finite impulse response filters (the output at each point depends only on a finite portion of the input). The filtered outputs are then down sampled by a factor of 2 in horizontal direction. This simply means that every second row of the resulting image is kept. These signals (sub images) are then each filtered by an identical filter pair in the vertical direction and down sampled by a factor of 2 again. The final result is a decomposition of the image into four subbands. An approximation subband, LL in Fig. 1(b), and details subbands - LH, HL and HH in Fig. 1(b). A two and three level decomposition can be seen in Fig. 1(c). and 1(d). In compression scenarios, one would discard most or all of the HL, LH and HH details and proceed to the second decomposition only using the LL subband as input.

2.2 Wavelets in face recognition so far

Li and Liu [6] used wavelet transformation prior to PCA in their work and reported a considerable performance improvement. Zhang *et al.* [7] applied Daubechies D8 wavelet transform to original face images, kept only the LL subband (similar to our experiment, as explained in Section 3) and used it as an input into the kernel associative memory neural network. They noticed a considerable improvement in recognition results. Ekenel and Sankur [8] made a multiresolution face recognition system using Daubechies 4 wavelet and PCA and ICA for classification. They noticed substantial improvements in performance for PCA and ICA in their expression changes and illumination changes experiments. None of those papers used CDF 9/7 wavelet for decomposition. These papers (and similar ones using DCT, e.g. [5]) inspired us to investigate the possibilities of using JPEG2000 compression scheme for face recognition.

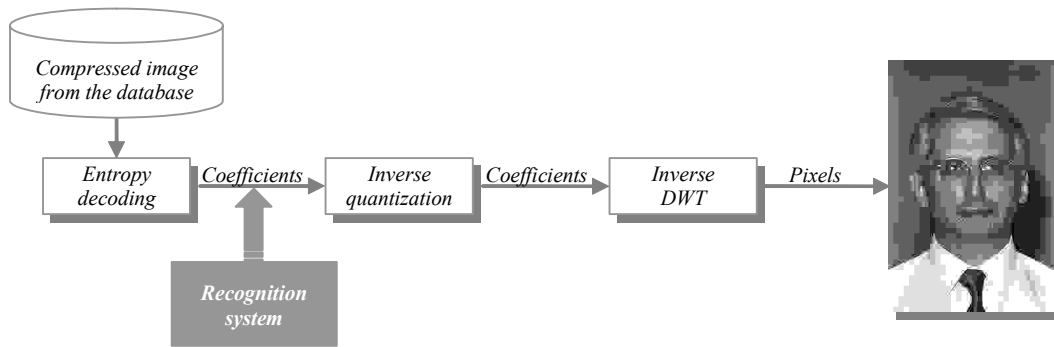


Fig. 3. Position of the classification (face recognition) system in the decompression chain in our experiments.

3. EXPERIMENTS

All recognition algorithms were implemented in Matlab following the description in cited papers. We developed our own instance of FERET performance evaluation in Matlab as well. Standard JPEG2000 coder implementation, available at www.jpeg.org, and a tool available at [19] were used for extracting DWT coefficients.

3.1 Image database

Our experiment was performed on a standard grey FERET data set [12], consisting of images of 1196 individuals taken under various conditions and at various points in time. Also, to achieve highly reproducible results, standard test sets were used, i.e. *fb* (different expression test), *fc* (different illumination), *dup1* (images taken anywhere between one minute and 1,031 days after the gallery image) and *dup2* (images taken at least 18 months after the gallery image was taken). Images were pre-processed in a standard way: rotated (so that the eyes are in the same position across all images), cropped to 128×128 pixels and then histogram equalized. These canonical images were then subjected to our experiment.

Prior to all experiments, we performed geometrical normalization of face images. This was done intentionally and enabled us to directly compare the results to previous studies. We are well aware that this approach somewhat violates the premise that we are working strictly in the compressed domain. We rely here on the fact that eventually algorithms that can localize eyes in compressed domain will be developed (an initial effort can be seen in [20]) and in our study we mimic this procedure prior to experimenting.

3.2 Pixel domain experiment

In our first experiment, PCA [9], ICA [10] and LDA [11] were performed on the original 128×128 images and the results for all FERET tests were noted. PCA is a subspace projection technique widely used for face recognition. It finds a set of representative projection vectors such that the projected samples retain most information about original samples. The most representative vectors are the eigenvectors corresponding to the largest eigenvalues of the covariance matrix. While PCA deals with variance (second-order statistics), ICA captures both second and higher-order statistics and projects the input data onto the basis vectors that are as

statistically independent as possible. We can state that ICA is a generalization of PCA. LDA, unlike PCA or ICA, uses the class information and finds a set of vectors that maximizes Fisher Discriminant Criterion. It simultaneously maximizes the between-class scatter while minimizing the within-class scatter in the projective feature vector space. While PCA and ICA can be called unsupervised learning techniques, LDA is supervised learning technique because it needs class information for each image in the training process.

After PCA, top 40% of eigenvectors were kept and that subspace (denoted by W in further text) was used for PCA recognition and also as the input to ICA and LDA algorithms (subspaces yielded by ICA and LDA will also be denoted by W). For algorithm training stage, 675 images of 225 subjects were used (those subjects for which there are exactly 3 images per person in the database). In testing stage, standard FERET gallery and probe sets (delivered with the database) were used. Algorithms were implemented with the most commonly used metrics, i.e. L1, L2 and COS. The results obtained this way with uncompressed images (pixel domain) are used as control results for the subsequent tests.

3.3 Compressed domain experiments

The second and third experiment (the compression experiments) are performed in exactly the same way as the first one, with the only difference being the fact that input to PCA, LDA and ICA are now DWT coefficients instead of image pixels. Fig. 3 shows where the recognition system is logically placed within the decompression procedure scheme. Fig. 4 shows the detailed experimental setup for the compression experiments.

The training stage is shown in the upper part of Fig. 4. The algorithms are trained with the same set of 675 images that are now compressed to a given number of bits and then partly decompressed, thus yielding the DWT coefficients. Those coefficients are distorted by compression. This experiment, and consequently this paper, actually investigates two things: first, the possibility of using CDF 9/7 coefficients in face recognition in general and second, the effect that DWT coefficient distortion introduced by compression has on recognition performance. In the testing stage (lower part of Fig. 4) the same compression and partial decompression procedure is repeated and DWT coefficients are used as input features to the recognition stage.

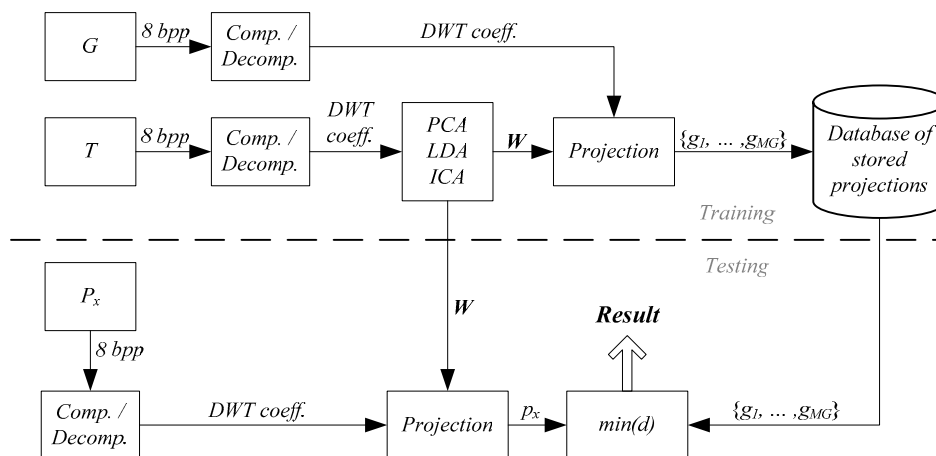


Fig. 4. Experimental setup for the compression experiments; MG is the number of images in the gallery set, p_x is the probe image to be recognized and W is the subspace that each algorithm calculates and d is the distance between a given gallery image and a probe image in projection space. In the "Projection" block, feature vectors (DWT coefficients rearranged into a vector) are mapped onto a subspace W such that projection $p = W^{-1}(X - x)$, where X is the feature vector and x is the mean of all feature vectors from the training set (for more details please refer to [9], [10], [11]).

The same number of input elements is used (128×128) since no coefficients are ignored or discarded. In the second test the DWT coefficients are the result of partly decompressing the images that were compressed to 1 bpp. The third test differs only in compression rate which is now 0.5 bpp.

4. RESULTS AND DISCUSSION

The results of our experiments can be seen in Table 1. Even at first glance it is obvious that most recognition algorithms experience no significant performance drop for all probe sets. The fact that this effect is persistent across all algorithms tells us that this is a property of the wavelet transform and not recognition methods. For the fb probe the results are very stable in both compression experiments (at 1 bpp and 0.5 bpp). Performance drop is less than 2% in most cases. Surprisingly, there are a few cases where recognition rate is higher at 0.5 bpp than at 1 bpp (e.g. PCA+L1). For the $dup1$ and $dup2$ probe sets this effect is not so emphasized, even though the differences are statistically insignificant or comparable with the results obtained by using original images. This is in some contradiction with the study reported in [13] where it was found that the recognition rate often goes slightly up with JPEG2000 compression (notice, those results are obtained using compression and then decompression, thus working in pixel domain only). This is probably due to the fact that we used a very simple feature extraction technique (i.e. we simply used all available coefficients, zero and nonzero) in this research and this fact only encourages us to do further research into efficient feature extraction techniques from standard JPEG2000 compression scenario. More intelligent feature vector formation, is clearly needed and will be part of our future work.

Nevertheless, JPEG2000 quantization and entropy coding obviously eliminate the DWT coefficients not essential for discrimination. Earlier studies concluded that information in low spatial frequency bands plays a dominant role in face recognition. Nastar et al. [21] have investigated the relationship between variations in facial

appearance and their deformation spectrum. They found that facial expressions and small occlusions affect the intensity manifold locally. Under frequency-based representation (such as wavelet transform), only high-frequency spectrum is affected. As for the fc probe set, most of the results for our two compression experiments are comparable to the ones obtained with original images. However, we noticed an improvement in performance for all PCA and LDA algorithms. The increase in recognition ratio at 0.5 bpp when compared to 1 bpp is even more obvious for this probe set. It is also interesting to compare these results with the ones reported in [13]. There, the fc test, after compression and decompression (i.e. in pixel domain), had the least improvement in recognition ratio. Again, this is encouraging for finding new ways of extracting feature vectors in JPEG2000 compressed domain (avoiding inverse DWT) thus improving performance for this difficult task. An optimal feature vector should improve current $dup1$ and $dup2$ (images taken later in time) results while retaining some of the advantages we noticed for the fc (different illumination) task.

5. CONCLUSIONS AND FURTHER WORK

In this study we showed that face recognition in compressed JPEG2000 domain is possible, since there are no significant performance drops. Face recognition algorithms can therefore be implemented directly into the JPEG2000 compressed domain without fear of deleterious effect on recognition rate. Such an implementation would save a considerable amount of computation time (due to avoiding the inverse DWT) and storage and bandwidth requirements (due to the fact that images could be compressed). One interesting result that needs to be emphasized here is the improvement in recognition rate for PCA and LDA algorithms for the fc probe set. This further justifies research into possible implementation of face recognition algorithms directly into JPEG2000 compressed domain, as it could (as a bonus benefit) also improve performance for different illumination task.

Algs.	Original images				DWT coefficients (JPEG2000 at 1 bpp)				DWT coefficients (JPEG2000 at 0.5 bpp)			
	<i>fb</i>	<i>fc</i>	<i>dup1</i>	<i>dup2</i>	<i>fb</i>	<i>fc</i>	<i>dup1</i>	<i>dup2</i>	<i>fb</i>	<i>fc</i>	<i>dup1</i>	<i>dup2</i>
PCA+L1	79.4	47.9	38.5	19.7	77.8	49.0	37.1	18.8	79.0	50.0	38.2	18.4
PCA+L2	76.4	11.3	34.9	14.1	75.0	19.6	32.4	8.5	75.1	19.6	33.0	9.8
PCA+COS	74.7	10.8	35.3	13.7	73.8	19.6	33.9	10.7	73.9	18.6	33.8	10.3
LDA+L1	72.9	11.9	35.3	16.7	72.3	18.6	34.6	15.0	72.6	19.6	35.2	15.0
LDA+L2	76.9	11.9	34.3	13.2	75.6	22.2	32.7	9.0	75.7	23.2	33.0	9.8
LDA+COS	75.4	11.3	35.6	12.8	74.1	21.6	34.1	10.3	74.6	21.1	34.2	10.3
ICA+L1	67.3	18.6	35.0	23.5	65.9	18.0	32.4	22.2	65.3	13.9	31.6	21.4
ICA+L2	76.8	47.9	36.2	23.5	75.7	45.4	33.7	23.5	75.5	46.4	33.2	22.7
ICA+COS	83.0	68.6	44.3	30.8	83.0	68.0	42.9	31.6	82.8	67.5	43.5	31.6

Table 1. Results of all the experiments. The numbers in the table represent rank 1 recognition rate percentages.

That should be backed up by testing using databases other than FERET since it is questionable how representative are the images in FERET of the entire population and how do any results obtained using FERET generalize to other databases.

Our further work will consist of finding a way to intelligently extract feature vectors within the standard JPEG2000 compression scheme (before inverse DWT) in order to improve recognition performance.

ACKNOWLEDGMENT

The work described in this paper was conducted under the research project: "Intelligent Image Features Extraction in Knowledge Discovery Systems" (036-0982560-1643), supported by the Ministry of Science, Education and Sports of the Republic of Croatia.

Portions of the research in this chapter use the FERET database of facial images collected under the FERET program. The authors would like to thank the FERET Technical Agent, the U.S. National Institute of Standards and Technology (NIST) for providing the FERET database.

REFERENCES

- [1] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, "Face Recognition in Still and Video Images: A Literature Survey", *ACM Comp. Surveys*, Vol. 35, pp. 399-458, 2003
- [2] C. Chistopoulos, A. Skodras, T. Ebrahimi, "The JPEG2000 Still Image Coding System: An Overview", *IEEE Trans. on Consumer Electronics*, Vol. 46, pp. 1103-1127, 2000
- [3] S. Eickeler, S. Müller, G. Rigoll, "High Quality Face Recognition in JPEG Compressed Images", *IEEE Int. Conf. on Image Processing (ICIP)*, Kobe, Japan, October 1999
- [4] S. Eickeler, S. Müller, G. Rigoll, "Recognition of JPEG Compressed Face Images Based on Statistical Methods", *Image and Vision Comp.*, Vol. 18, No. 4, pp. 279-287, 2000
- [5] W. Chen, M. J. Er, S. Wu, "PCA and LDA in DCT Domain", *Pattern Recognition Letters*, Vol. 26, No. 15, pp. 2474-2482, 2005
- [6] B. Li, Y. Liu, "When Eigenfaces are combined with Wavelets", *Knowledge-Based Systems*, Vol. 15, No. 5-6, pp. 343-347, 2002
- [7] B-L. Zhang, H. Zhang, S. S. Ge, "Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory", *IEEE Trans. on Neural Networks*, Vol. 15, No. 1, pp. 166-177, 2004
- [8] H.K. Ekenel, B. Sankur, "Multiresolution Face Recognition", *Image and Vision Computing*, Vol. 23, No. 5, pp. 469-477, 2005
- [9] M. Turk, A. Pentland, "Eigenfaces for Recognition", *J. of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991
- [10] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, "Face Recognition by Independent Component Analysis", *IEEE Trans. on Neural Net.*, Vol. 13, No. 6, pp. 1450-1464, 2002
- [11] P. Belhumeur, J. Hespanha, D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", *Proc. of the 4th European Conf. on Computer Vision, ECCV'96*, Vol. 1, 1996, Cambridge, UK, pp. 45-58
- [12] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1090-1104, October 2000
- [13] K. Delac, M. Grgic, S. Grgic, "Effects of JPEG and JPEG2000 Compression on Face Recognition", *LNCS*, Vol. 3687, pp. 136-145, 2005
- [14] S.L. Wijaya, M. Savvides, B.V.K. Vijaya Kumar, "Illumination-tolerant face verification of low-bit-rate JPEG2000 wavelet images with advanced correlation filters for handheld devices", *Appl. Opt.*, Vol. 44, pp. 655-665, 2005
- [15] D. Valentin, "Face-space Models of Face Recognition", *Computational, Geometric, and Process Perspectives on Facial Cognition: Contexts and Challenges*, 1999
- [16] I. Daubechies, "Ten Lectures on Wavelets", CBMS, *Society for Industrial and Applied Mathematics*, 1992
- [17] A. Cohen, I. Daubechies, J.C. Feauveau, "Biorthogonal Bases of Compactly Supported Wavelets", *Comm. on Pure and Applied Math.*, Vol. 45, No. 5, pp. 485-560, May 1992
- [18] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, "Image Coding Using Wavelet Transform", *IEEE Trans. on Image Processing*, Vol. 1, No. 2, pp. 205-220, April 1992
- [19] N. Sprljan, MATLAB Wavelet Toolbox v1.02. Available: <http://www.sprljan.com/nikola/matlab/wavelet.html>, 2007
- [20] P. Fonseca, J. Nasvadba, "Face Detection in the Compressed Domain", *Proc. of the 2002 IEEE International Conference on Image Processing*, Singapore, 24-27 October 2004, Vol. 3, pp. 2015-2018
- [21] C. Nastar, N. Ayach, "Frequency-based Nonrigid Motion Analysis", *IEEE Trans. on PAMI*, Vol. 18, pp. 1067-1079, 1996

