Image and Vision Computing 27 (2009) 1108–1120

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Face recognition in JPEG and 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

ARTICLE INFO

Article history:
Received 28 January 2008
Accepted 12 October 2008

Keywords:
Face Recognition
Compressed Domain
JPEG
JPEG2000
DCT
DWT

ABSTRACT

In this paper we investigate the potential of performing face recognition in JPEG and JPEG2000 compressed domain. This is achieved by avoiding full decompression and using transform coefficients as input to face recognition algorithms. We propose a new comparison methodology and by employing it show that face recognition can efficiently be implemented directly into compressed domain. In the first part of our experiment we use all the available transform coefficients and show that recognition rates are comparable and in some cases even higher than recognition rates obtained by using pixels from uncompressed images (standard face recognition approach). In the second part, we propose an effective coefficient preselection method (applicable both in JPEG and JPEG2000 compressed domain). Our results show that by using the proposed method, recognition rates can be significantly improved while additionally reducing computational time. Finally, we propose what a hypothetical compressed domain face recognition system should look like.

1. Introduction

Automatic face recognition (AFR) is currently, along with other biometric methods, one of the most vividly researched areas in computer science [1–4]. The problems to be solved in this area are challenging and therefore attractive to a wide range of researchers with different backgrounds. AFR is difficult primarily because of variations that images of human faces undergo in any real life scenario. Some examples are variations in imaging conditions (lighting and viewpoint changes induced by body movement) and differences induced by effects such as ageing, facial expressions and occlusions. Researchers in this field have up to recently used uncompressed high-resolution still images in their research. However, influenced by other research areas (such as image indexing and retrieval) and boosted by real life implementation needs (storage requirements and computational speed), they slowly started to consider using compressed images in AFR. Compression was recognized as an important issue and is an actively researched area in other biometric approaches as well. Most recent efforts have been made in iris recognition [5,6] and fingerprint recognition [7,8]. Apart from trying to deploy standard compression methods in recognition, researchers even develop a special purpose compression algorithms, e.g. a recent low bit-rate compression of face images [9].

There are several important reasons for introducing image compression in AFR. The most obvious one is that e-passports (see also recently released Face Recognition Format for Data Interchange [10,11]) will use face images as one of the three identifiers (with fingerprints and iris scans being the other two). To be able to do so, a face image will have to be stored on a limited capacity chip and compression thus becomes inevitable [12]. Another reason is the need to store large number of images for any law-enforcement purposes. When storing hundreds of thousands images, compression becomes an important issue. Furthermore, image acquisition equipment also often delivers already compressed images at its output. Working directly in compressed domain has its reasons as well. For the stored (or outputted) image of the unknown individual to be recognized/identified it has to be compared to one or more images of known individuals. Decompressing all those images (including the unknown one) is very computationally intensive. Avoiding full decompression from this point of view seems quite attractive, and advantages of similar approaches have already been shown in [13].

The general focus of this paper is AFR from still images, although the proposed methodology can be translated into video-based recognition. More precisely, the main focus of this paper is *face recognition in compressed domain*, with JPEG [14] and JPEG2000 [15] being the obvious choice for compression standards. We feel that common image compression standards, such as JPEG and JPEG2000, have the highest potential for actual usage in real life, since the image will always have to be decompressed and presented to a human at some point in the process. From that perspective it seems reasonable to use standardized and commonly implemented compression format. Working directly in compressed

* Corresponding author. Tel.: +385 1 6129 851; fax: +385 1 6129 568/+385 1 6129 717.
E-mail addresses: kdelac@ieee.org (K. Delac), mgrgic@ieee.org (M. Grgic), sgrgic@ieee.org (S. Grgic).

0262-8856/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2008.10.007

domain solves both the storage requirements problem (images in compressed form use less storage space) and computational speed problem (being able to partly decompress the images speeds up the whole recognition process).

Effects of JPEG and JPEG2000 compression on face recognition performance have been extensively researched in the past. A small experiment using JPEG can be found in [16] and more extensive experiments using both JPEG and JPEG2000 can be found in [12,17,18]. All those studies mainly agree that *compression does not deteriorate recognition rate significantly* (up to a certain compression ratio). This conclusion can be considered as basic assumption needed to proceed to compressed domain recognition. Moreover, it was found in [17] and [18] that in some cases (at certain compression ratios) the recognition rate is even slightly improved when using compressed images. All these aforementioned experiments were performed in pixel (uncompressed) domain, such that images were first compressed at a given compression ratio and then fully decompressed prior to recognizing. In all the mentioned experiments only probe images were compressed, and training and gallery images were uncompressed. This methodology poses certain problems when trying to compare results obtained in compressed domain to those experiments and we shall bypass these problems with our proposed methodology described in section 3.

It is also worthwhile to define the *compressed domain*, as this is a basic term in our research. To perform image compression, pixel values are firstly transformed into transform coefficients, transform coefficients are quantized (discretized) and quantized coefficients are entropy coded. Compressed domain means that *transform coefficients* are used as input to face recognition methods instead of pixel values. JPEG uses Discrete Cosine Transform (DCT) and JPEG2000 uses Discrete Wavelet Transform (DWT), so in compressed domain DCT or DWT coefficients will be used for face recognition. Formally, compressed domain is any point in the compression/decompression procedure after transform and before inverse transform (Fig. 1). We shall refer to data before compression and after decompression as data in the *pixel domain*. Images that were not compressed at all will be referred to as *uncompressed images* or *original images*, and working with them will be considered as working in pixel domain as well. Inverse (discrete cosine or wavelet) transform is computationally the most intensive part of decompression process and avoiding it in face recognition systems is therefore desirable. Coefficients taken before the inverse transform will be distorted by quantization in the compression process and the influence of this distortion on face recognition accuracy needs to be examined. This is an important aspect and contribution of this paper and this issue, along with our proposed coefficient preselection method, is thoroughly discussed in sections 4 and 5.

In this paper we shall use two well known face recognition methods: Principal Component Analysis (PCA) [19] and Independent Component Analysis (ICA) [20]. By compressing images from standard FERET database [21,22] and using transform coefficients directly from the compressed domain we shall prove that face recognition can be performed directly in compressed domain. Furthermore, we shall not only use all available coefficients but also propose a coefficients preselection scheme that will speed up the whole recognition process and in some cases drastically improve recognition accuracy. All our conclusions will be supported by extensive experiments and McNemar's hypothesis test for statistical significance of the observed results.

The rest of this paper is organized as follows: in section 2 a review of recent related work is given; in section 3 we present our methodology used in experiments; in section 4 we describe experimental setup and present recognition results when all available transform coefficients are used instead of pixels; in section 5 a coefficient preselection method is proposed and detailed experimental analysis of the proposed approach is presented; section 3 concludes the paper.

2. Related work

Lately there has been a lot research done on the subject on various image analysis methods implemented directly in compressed domain [13]. We shall here give a brief overview of some papers that are related to the subject of this paper. This section is divided into three subsections. The first one deals with papers that use DCT and/or JPEG in their recognition process, while the second one deals with papers that use DWT and/or JPEG2000. In the third subsection we summarize the described approaches.

2.1. DCT

One of the first works in the area of face recognition in compressed domain can be found in [23]. The authors use binary keys derived from DCT coefficients in JPEG compression procedure. Even though they use standard JPEG compression, there is no mention of the exact compression ratio used and there is no analysis on how compression affects the recognition results. They used self made database of faces. Seales et al. in [24] give a detailed mathematical analysis of PCA (as a pattern recognition tool) and JPEG compression and suggest how the two can be combined into a unique system working in compressed domain. The results are shown as a function of quality factor and the overall conclusion is that there

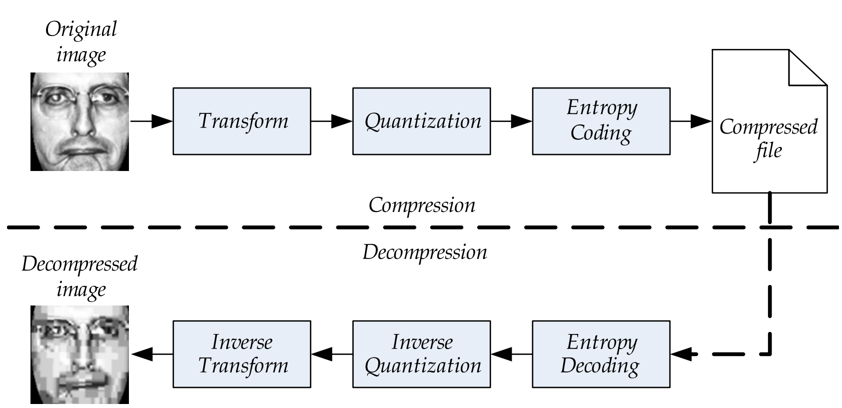

Fig. 1. Transform coding.

are substantial savings in computation time when working in compressed domain.

Eickeler et al. [25,26] use DCT coefficients as input into Hidden Markov Model (HMM) used for classification. Standard JPEG procedure is used where the compressed image is entropy decoded and inversly quantized. The features are then extracted from the yielded dequantized coefficients using the first 15 coefficients from each 8×8 block. The final matrix of size $15 \times n$ (where n is the number of block in an image) is used as a feature that is an input to HMM. The proposed system is tested on a database of faces of 40 individuals. A slight increase in performance (5.5% higher recognition rate) was detected for images compressed approximately to 1 bpp in comparison with uncompressed images.

Hafed and Levine performed DCT over the whole image (not related to JPEG) in [27] and kept 49 coefficients to be used as input into standard PCA-based recognition system. They detected a 7% recognition improvement compared to original (uncompressed) images. The proposed method was tested on a couple of small databases and the results were given in table form for rank one and as Cumulative Match Curves (CMS) [22] for higher ranks. Tyahyadi et al. [28] perform DCT over 8×8 blocks and use energy histogram analysis over the calculated coefficients. Feature vectors are composed from the those histograms and Euclidean distance is calculated between each of them. They test their system on a database of 15 individuals (165 images in total) and detect a performance increase of about 10% compared to standard PCA with uncompressed images.

Chen et al. [29] use the FERET database of face images and perform DCT on 8×8 blocks of pixels after which they quantize the coefficients with a standard JPEG quantization matrix (for the luminant component). They use the quantized coefficients (all of them) as input into the PCA and LDA recognition algorithms. The authors show that recognition results in compressed domain are the same as with uncompressed images in pixel domain. Furthermore, instead of using all the DCT coefficients, the authors also show that using only the topmost 20 coefficients (in a JPEG zigzag analysis sense) will also not deteriorate the results significantly.

2.2. DWT

Sabharwal and Curtis [30] analyze the images with Daubechies 2 wavelet filter and use the extracted coefficients as input to PCA. They perform one, two and three decompositions and show that recognition resutls increase compared to PCA with uncompressed images. Effect of increased recognition performance with increasing number of decompositions is also observed, although the results in this case are only marginaly better (about 2% higher recognition rate). A small custom database of images was used.

In [31] two wavelet decompositions were used, but the second decomposition is not performed on the approximation band of the first decomposition only, but on each band individually. Thus, two wavelet decompositions yielded 16 band altogether and those coefficients were used as features for classification. Battacharyya distance was used as a classifier and the results were compared to a standard PCA+L2 recognition algorithm. FERET database of images was used. The results show that wavelet analysis improves overall results. The same conclusion, using standard two decompositions, was reached in [32] as well. Li et al. [33] use wavelet coefficients as input to PCA and also get consistently higher recognition rates compared to standard PCA.

Zhang et al. [34] use two decompositions (Daubechies wavelet) and keep only the coefficients from the approximation band. Those coefficients are used as input to a neural network used as a classifier. Several image databases (including FERET) are used and the results are compared to standard PCA on the same images. The authors report recognition rate increase in all experiments.

Ekenel and Sankur [35] used Daubechies 4 wavelet and PCA and ICA as recognition algorithms. They tried to find the wavelet subbands that are least sensitive to illumination and expression changes. They combine images from several databases which makes the results difficult to compare to. However, this study is performed in a very scientifically strict manner since the same recognition method is used once with uncompressed pixels as input (what we so far referred to as standard PCA method) and once with DWT coefficients as input. In the experiment with images of different expressions no significant difference in recognition results using uncompressed images and DWT coefficients was observed. In the experiment with images with different illumination conditions a considerable improvement was observed when DWT coefficients were used instead of pixels (over 20% higher recognition rate for all tested methods).

2.3. Analysis

We shall now try to give a short summary and synthesis of the papers reviewed above. First of all, most of them are not directly linked to JPEG or JPEG2000 standards. The ones that are linked to JPEG or JPEG2000 do not give their analysis based on compression ratio. Furthermore, authors often do not try to preselect the coefficients but simply either give all of them as input to recognition algorithms, or take only the first couple from each block (in a JPEG-like scheme). Even though there are some papers that address the problem of recognition in compressed JPEG domain, compressed JPEG2000 domain remains unexplored. Initial study of face recognition in JPEG2000 compressed domain was performed in [36] and preliminary results shown there will be largely expanded in this paper.

Second conclusion that can be drawn is that there is a general lack of a systematic methodology for comparing the results in compressed domain to those in pixel domain. Testing statistical significance of the observed performance differences is hardly ever done. Using larger publicly available databases with accompanying test sets (such as FERET) is rarely the practice, even though this is one of the key factors of performing reproducible research.

In this paper we shall try to address almost all of the problems mentioned above. Our starting point was the work done in [29,35] and [36]. We shall work directly in JPEG and JPEG2000 compressed domain using standard codec implementations. We shall test two well known recognition algorithms on the FERET database following the accompanying protocol for identification scenario [22]. By proposing and then using the proposed measurement and comparison methodology we shall show that face recognition can efficiently be implemented directly into compressed domain, thus reducing computational time and storage requirements. Finally, we propose a simple yet effective coefficient preselection method (applicable both in JPEG and JPEG2000 compressed domain) and show that recognition rates can this way be significantly improved while additionally reducing computational time.

3. Methods

3.1. Recognition algorithms

In our first experiments, PCA [19] and ICA [20] were performed on the original (uncompressed) 128×128 images and the results for all FERET tests were noted (Table 1–4, the “Orig.” column). PCA [19] is a subspace projection technique widely used for face recognition. Given an s-dimensional vector representation of each face in a training set of images, PCA tends to find a t-dimensional subspace whose basis vectors correspond to the maximum variance direction in the original image space. This new subspace is

Table 1
Results (rank 1 recognition rates) for all probe sets obtained by using all 16384 DCT coefficients per image as input to PCA and ICA (JPEG compressed domain)

JPEG - DCT$_{16384}$			*Orig.*	*1 bpp*	*0.5 bpp*	*0.3 bpp*	*0.2 bpp*
fb	PCA	RR	79.4	80.3	76.7	74.9	74.4
		McN	-	○	✗	✗	✗
	ICA	RR	83.0	83.2	84.1	82.3	82.5
		McN	-	○	○	○	○
fc	PCA	RR	47.9	62.9	61.3	51.0	33.0
		McN	-	✓	✓	○	✗
	ICA	RR	68.6	71.6	71.1	68.0	57.2
		McN	-	○	○	○	✗
dup1	PCA	RR	38.5	40.3	34.3	33.4	36.0
		McN	-	✓	○	○	✗
	ICA	RR	44.3	46.7	42.5	43.9	39.1
		McN	-	✓	✗	○	✗
dup2	PCA	RR	19.7	21.4	22.2	20.5	17.5
		McN	-	○	○	○	○
	ICA	RR	30.8	34.2	30.8	32.1	23.9
		McN	-	✓	○	○	✗

Table 2
Results (rank 1 recognition rates) for all probe sets obtained by using all 16384 DWT coefficients per image as input to PCA and ICA (JPEG2000 compressed domain)

JPEG2000 - DWT$_{16384}$			*Orig.*	*1 bpp*	*0.5 bpp*	*0.3 bpp*	*0.2 bpp*
fb	PCA	RR	79.4	77.8	79.0	79.5	80.6
		McN	-	✗	○	○	○
	ICA	RR	83.0	83.0	82.8	83.8	83.6
		McN	-	○	○	○	○
fc	PCA	RR	47.9	49.0	50.0	52.6	53.6
		McN	-	○	✓	✓	✓
	ICA	RR	68.6	68.0	67.5	65.5	57.5
		McN	-	○	○	○	✗
dup1	PCA	RR	38.5	37.1	38.2	38.1	36.6
		McN	-	○	○	○	✗
	ICA	RR	44.3	42.9	43.5	43.5	38.9
		McN	-	○	○	○	✗
dup2	PCA	RR	19.7	18.8	18.4	17.1	16.2
		McN	-	○	○	○	✗
	ICA	RR	30.8	31.6	31.6	32.1	25.2
		McN	-	○	○	○	✗

Table 3
Results (rank 1 recognition rates) for all probe sets obtained by using the preselected 512 DCT coefficients per image as input to PCA and ICA (JPEG compressed domain)

JPEG - DCT$_{512}$			*Orig.*	*1 bpp*	*0.5 bpp*	*0.3 bpp*	*0.2 bpp*
fb	PCA	RR	79.4	82.5	79.1	74.9	76.5
		McN	-	✓	○	✗	✗
	ICA	RR	83.0	84.4	84.1	81.9	80.0
		McN	-	✓	○	○	✗
fc	PCA	RR	47.9	77.3	73.7	67.0	49.5
		McN	-	✓	✓	✓	○
	ICA	RR	68.6	73.2	69.6	68.0	48.5
		McN	-	✓	○	○	✗
dup1	PCA	RR	38.5	39.1	33.9	34.1	37.1
		McN	-	○	✗	✗	○
	ICA	RR	44.3	43.5	41.6	41.1	40.4
		McN	-	○	✗	✗	✗
dup2	PCA	RR	19.7	22.2	20.9	21.8	17.9
		McN	-	○	○	○	○
	ICA	RR	30.8	35.0	32.1	29.9	29.5
		McN	-	✓	○	○	○

normally lower dimensional ($t \ll s$). In other words, 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 (for further details please refer to www.face-rec.org/algorithms/#PCA).

Table 4
Results (rank 1 recognition rates) for all probe sets obtained by using the preselected 512 DWT coefficients per image as input to PCA and ICA (JPEG2000 compressed domain)

JPEG2000 - DWT$_{512}$			*Orig.*	*1 bpp*	*0.5 bpp*	*0.3 bpp*	*0.2 bpp*
fb	PCA	RR	79.4	81.7	81.4	82.0	81.4
		McN	-	✓	✓	✓	✓
	ICA	RR	83.0	83.7	83.5	84.4	83.3
		McN	-	○	○	✓	✗
fc	PCA	RR	47.9	66.5	65.5	64.9	60.8
		McN	-	✓	✓	✓	✓
	ICA	RR	68.6	67.5	69.1	62.9	56.2
		McN	-	○	○	✗	✗
dup1	PCA	RR	38.5	38.4	38.1	38.1	36.7
		McN	-	○	○	○	○
	ICA	RR	44.3	44.6	45.3	44.2	38.5
		McN	-	○	○	○	✗
dup2	PCA	RR	19.7	18.8	19.2	19.2	17.5
		McN	-	○	○	○	○
	ICA	RR	30.8	35.0	35.9	34.6	27.8
		McN	-	✓	✓	○	○

While PCA deals only with second-order statistics (variance), ICA [20] captures both second and higher-order statistics and projects the input data onto the basis vectors that are as statistically independent as possible. Our implementation of ICA uses the INFOMAX algorithm proposed by Bell and Sejnowski and used in [20]. PCA is performed to reduce dimensionality prior to ICA. ICA Architecture 2 is used as it gives consistently higher recognition rates in the identification scenario [37]. For further details on ICA please refer to www.face-rec.org/algorithms/#ICA.

After performing PCA, top 40% of eigenvectors were kept and that subspace was used for PCA recognition and also as the input to ICA. Both final subspaces in which recognition will be performed (the one yielded by PCA and the one yielded by ICA) will be denoted by $\boldsymbol{W}$ in the rest of this paper. For algorithm training stage, 675 images of 225 subjects were used. This training set will be denoted as T, with the number of images $M = 675$. In testing stage, standard FERET gallery and probe sets (delivered with the database) were used. The results obtained this way with uncompressed images are used as control results for the subsequent tests (please see [37] for a more detailed description of a general pixel experiment).

As a distance metric in the calculated subspaces we decided to use the L1 (City block distance) for PCA (the whole recognition algorithm will be denoted as PCA in further text) and Cosine angle for ICA (the whole recognition algorithm will be denoted as ICA in further text). The aforementioned metrics performed best in many of our previous studies [17,37] and that was the main reason for choosing them.

3.2. FERET database

Our experiment was performed on a standard grey FERET data set [22], consisting of images of 1196 individuals taken under various conditions and at various points in time. To achieve highly reproducible results, standard test sets were used: *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 geometrically normalized (preprocessed) 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 experiments.

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 rely here on the fact that eventually algorithms that can localize eyes and thus geometrically normalize face images in compressed domain will be developed (an initial effort can be seen in [38]).

3.3. Compression

All images used in experiments were compressed according to JPEG and JPEG2000 compression standards, with various compression ratios (bitrate, bpp): 1, 0.5, 0.3 and 0.2 bpp. To compress images using JPEG, the Independent JPEG Group's JPEG software packet (JPEG6b32) was used [39]. To yield various bitrates, quality parameter was iteratively set until the desired bitrate was achieved. To compress images using JPEG2000 standard, a JJ2000 V4.1 (up to date with Part 1 of the JPEG2000 standard) was used [39]. An interface to JPEG and JPEG2000 implementations was built in Matlab, which served as a platform for experimenting and results analysis. Compression was done on the preprocessed images.

3.4. Performance measurement methodology

3.4.1. Recognition rate

The main tool for measuring recognition accuracy in this paper will be the recognition rate. Its definition can generally be formulated as follows. For each probe image $p_{i,j} \in P$, sort all the gallery images by decreasing similarity, yielding a list $L = \{L_1, L_2, \ldots, L_K\}$, where K is the total number of subjects in the gallery (assuming that there is one image per subject, K also becomes the number of images and the size of the gallery), i.e. K is the number of images in G. L_1 is the gallery image most similar to the given probe image (according to the algorithm), L_2 is the next closest match and expanding this to L_k being the kth closest gallery match.

In this case, if L_1 (labeled as the closest gallery match to the given probe image) is really the correct answer (checked in the ground truth information) we say that the algorithm *correctly recognized the probe image*. In other words, the algorithm successfully recognizes a probe image if the probe image and the top ranked gallery image in L are of the same subject. This is called *rank 1 recognition rate* (because we are using only the *top ranked* gallery image) and can be formally defined over the whole set of probe images P as follows: let R_1 denote the number of correctly recognized probe images in L at $k = 1$ and $|P|$ be the probe set size, then $rank_1 = R_1/|P|$. A usual way to report rank 1 performance is to give it in a form of percentage, for example we say that some algorithm has 86% rank 1 recognition rate (RR) on a given gallery and probe set. Rank 1 result will be reported in tables in the following sections.

3.4.2. Normalized recognition rate

To be able to compare the results obtained in pixel domain (using uncompressed images) to those in the compressed domain, we propose a new performance measure - *Normalized Recognition Rate* (NRR). NRR shall be used for rank 1 results. The recognition rate obtained at a certain compression rate (bpp) is normalized to recognition rate in pixel domain for the same recognition method. The main presumption here is that the two experiments are completely the same (same training, gallery and probe sets). Let RR_{pixel} be the rank 1 recognition rate obtained in pixel domain and let RR_{coeff} be the rank 1 recognition rate obtained in the compressed domain. The NRR for one compression rate is then given by $NRR = RR_{coeff}/RR_{pixel}$. Expanding the same line of thought at multiple compression rates will yield a graph that shows how NRR depends on compression rate. When recognition rate for a given algorithm is the same in pixel domain and in compressed domain, $NRR = 1$. When recognition rate is higher in compressed domain than in pixel domain, $NRR > 1$, and $NRR < 1$ in the opposite case. The main advantage of this approach is that it forces the use of exactly the same experimental setup and exactly the same recognition algorithm in comparisons.

3.4.3. Statistical significance

To further support any conclusions made on algorithm performance, we shall use a statistical significance test called McNemar's hypothesis test [40]. This test was previously shown to be the most appropriate one for face recognition experiments [37,41,42] and an interested reader is referred to those papers for a more detailed analysis of the test. The cutoff threshold for the p-value in our experiments was set to 0.05 (5%).

4. Transform coefficients as features

The obvious way to transport face recognition into the compressed domain is to stop the decompression procedure before the inverse transformation and to use transform coefficients as input to recognition algorithms. A block scheme of a general transform decoding procedure is shown in Fig. 2. In our experiments, we avoided the inverse quantization as well as the inverse transformation. The reason for this is that we found no performance improvement or degradation when using the inversely transformed coefficients. By avoiding the inverse quantization some additional computation time was saved.

In the experiment described in this section we shall use all the coefficients available after the entropy decoding. In case of our normalized 128×128 sized images, this means that we use all 16384 coefficients per image (concatenated in a 1×16384 vector) as input to PCA and ICA.

4.1. Experimental setup

The block-scheme of the experimental setup can be seen in Fig. 3. This is an exact description of the procedure we performed in order to do the experiment, but it is not a real life implementation scheme. The difference is that, since we worked with uncompressed FERET database images, we first had to compress them and then partially decompress them to extract the transform coefficients (the "*Comp./Decomp.*" block in Fig. 3). In a real life scenario we would already have compressed images and would only have to partially decompress them to get the coefficients.

In the training stage (upper part of Fig. 3), the training set of images (T) is used to calculate the PCA and ICA subspaces at a given compression ratio. The algorithms are retrained for each compression ratio. Then, all gallery images (G) are projected onto those subspaces yielding a set of projections $\{g_1, \ldots, g_{MG}\}$, where MG is the total number of gallery images. These projections will later be used in the testing stage (lower part of Fig. 3), where they will be compared to the projection of probe (unknown) image. The unknown image P_x, which is a part of a set of unknown images for a particular experiment, is compressed and partially decompressed. The obtained coefficients are projected onto the same subspace as the gallery images (PCA or ICA), resulting in projection p_x. This projection is then compared to all the gallery projections in the "*min(d)*" module (named so since the minimum distance in feature space $\boldsymbol{W}$ represents the closest match). The closest match is labeled as the correct result.

4.2. Results and analysis

The results of this experiment can be seen in Tables 1 and 2. The "Orig." column of Tables 1 and 2 represents the RR obtained using uncompressed (original) images. The RR obtained with transform coefficients (i.e. in compressed domain) is presented in columns

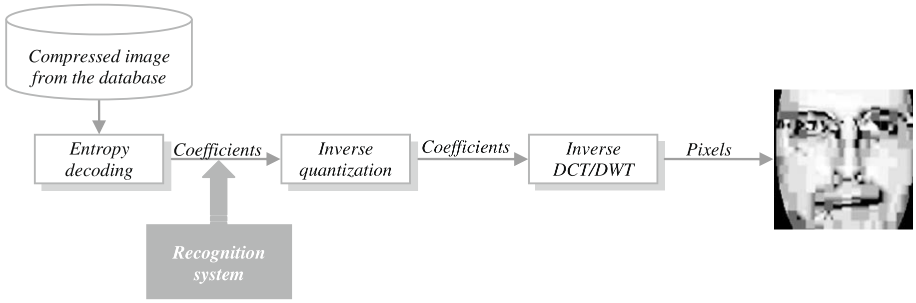

Fig. 2. Decompression procedure block scheme.

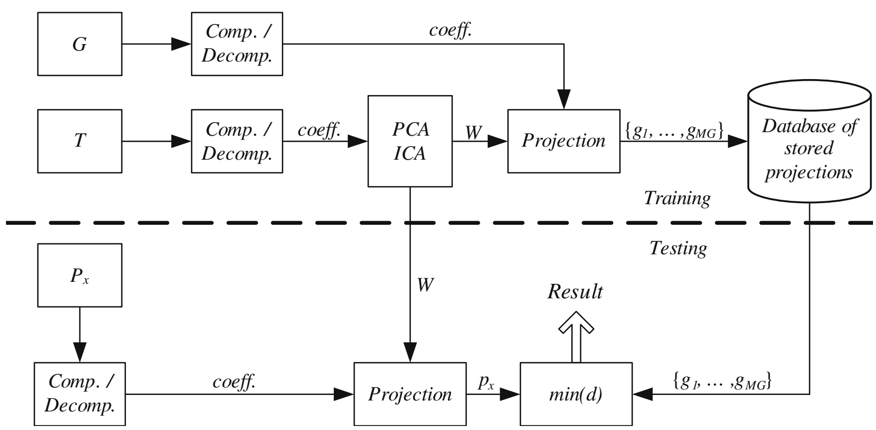

Fig. 3. Experimental setup.

that have compression ratios (expressed as number of bits per pixels, e.g. 1 bpp = 8:1 compression rate) as headers. Each table is divided into four segments corresponding to the four probe sets in FERET database. The "○" in the "McN" row means that there is no statistically significant performance difference (in terms of rank 1 recognition rate) between pixel and compressed domain; the "✘" means that the results in compressed domain are significantly worse in compressed domain than in pixel domain and "✔" means that the recognition rate in compressed domain in significantly better than in pixel domain.

By examining the results in Tables 1 and 2 we can generally conclude that the *use of transform coefficients as input to recognition algorithms does not significantly deteriorate recognition rates*. We shall further give a detailed analysis of the results.

4.2.1. Analysis by probe sets

For the *fb* probe set and DCT coefficients (Table 1) both recognition methods are quite stable. PCA shows slight, although statistically insignificant, improvement in recognition rate (RR) at 1 bpp, but experiences a more severe drop at higher compression rates. Even though the drops at 0.5, 0.3 and 0.2 bpp are statistically significant the absolute difference in RR for those cases is at best 5%. The RR at 0.5 bpp for instance is only 2.7% lower than with the uncompressed images, which is tolerable. Situation with ICA is somewhat different, as it shows a slight improvement in RR for 1 and 0.5 bpp, and also a slight drop for 0.3 and 0.2 bpp, both of which are statistically insignificant. Absolute differences compared to uncompressed image are approximately ±1%. When using DWT coefficients for the same probe set (Table 2) we see an interesting inversion for PCA. For DCT the highest RR was achieved at 1 bpp and for DWT we now see the lowest RR at 1 bpp. The absolute difference is only 1.6% but is statistically significant. Unlike the previous example with DCT, the RR at other (higher) compression rates with DWT is now statistically not significantly different compared to uncompressed images. We also observed an interesting increase of RR with the increase of compression rates.

This is probably due to JPEG2000 coefficients discarding (and coefficient precision degrading) during the compression procedure. Obviously, the coder efficiently discards the information that normally prevents PCA to correctly recognize some images. Since images in the *fb* probe set differ from gallery images only in facial expression, which is by its nature high frequency information, it is only natural that discarding exactly that information during compression helps PCA to match such images more efficiently. On the other hand, the coefficients that carry the information of a subject's identity are obviously neither discarded nor is their precision degraded. Similar conclusion stands for ICA as well, as it also shows a slight increase in RR at higher compression rates.

For the *fc* probe set there seem to be quite a few significant improvements in RR. In case of DCT coefficients and PCA (Table 1) there is a 15% improvement at 1 bpp and 13.4% at 0.5 bpp. Both improvements are statistically significant. The RR at 0.3 bpp is only a few percentages higher than with uncompressed images and then it drops significantly at 0.2 bpp. Slight improvement can be seen for ICA as well at 1 bpp and 0.5 bpp, and then a slight drop at 0.3 bpp (all three results are not significantly different from the result obtained with uncompressed images). Significantly lower RR is observed only at 0.2 bpp. In case of DWT coefficients (Table 2), PCA exhibits significant improvements (at 0.5, 0.3 and 0.2 bpp), with a constant increase of RR with the increase of compression rates. Similar effect was observed for the *fb* probe set as well. However, the improvements are in absolute far lower than with DCT coefficients. ICA performs relatively stable with no significant changes in RR up to 0.2 bpp. This probe set differs from the gallery in illumination conditions. Most of the differences arise from changing the overall illumination of the image, rather that the direction of the light source. Thus, most of the images in this set appear darker or lighter than the corresponding images in the gallery. When such images are compressed, the effect of "smearing" the image (in JPEG keeping the DC coefficients intact) decreases the differences between the images of the subject caused by such illumination changes. This is why there is a significant improvement in RR with DCT coefficients, while the same effect is less obvious with DWT coefficients. For those images that do have different direction of illumination (which is in its nature a low frequency information), discarding more and more coefficients or degrading their precision (thus making them more similar to corresponding coefficients in the gallery images) help the algorithms identify subjects from such images more efficiently. Once too many coefficients are discarded or degraded, the RR drops drastically. In case of PCA and DWT coefficients this drop occurs below 0.2 bpp.

The general reason for such a drop is that the remaining coefficients lost the information about the subject's identity. NRR curves for the *fc* probe set can be seen in Fig. 4. The curves marked "JPEG" and "JPEG2000" depict results obtained in an experimental setup as described in [17] and [18], where the images were first compressed to a certain compression ratio and then uncompressed prior to using them in recognition (pixel domain). Those results show how NRR changes at various compression ratios when image degradation caused by compression artifacts is introduced. The results of the experiment described in this section are marked "DCT (16384)" and "DWT (16384)", as all 16384 available DCT or DWT transform coefficients were used. Similarly, "DCT (512)" and "DWT (512)" depict the results of the experiment in section 5 where only a subset of 512 coefficients will be used.

It can clearly be seen from Fig. 4 that for the DCT coefficients, i.e. the JPEG compression (top two graphs in Fig. 4), the curves "DCT (16384)" remain above 1.0 for all compression rates up to 0.3 bpp. This means that the recognition rates are higher than with uncompressed images. Compared to the "JPEG" curve, the use of DCT coefficients instead of pixels obtained after decompression obviously improves performance, as the "DCT (16384)" curve is constantly above the "JPEG" curve. Similar effect can be noticed for the DWT coefficients, i.e. the JPEG2000 compression (bottom two graphs in Fig. 4), although less emphasized.

There are significant improvements in RR for the *dup1* and *dup2* sets (the images taken at different points in time) for DCT coefficients at 1 bpp, Tables 1 and 2. At higher compression rates the RR drops, in some cases even statistically significant. The information necessary for identification seems to be lost when the images are compressed below 1 bpp. This would be important when deciding at which compression rate any hypothetical face recognition system in compressed domain should work. For the DWT coefficients the differences between RR using uncompressed images and using coefficients are even lower, but become significantly worse only at 0.2 bpp. We can conclude that when DWT coefficients are used, the RR is more stable.

4.2.2. Analysis in feature space

We shall now try to illustrate what happens in the feature space (***W***) of PCA when DCT (or DWT) coefficients are used instead of pixels. To explore that, we performed a small experiment using 50 randomly chosen images of *different subjects*, calculated the feature space ***W*** from them, retained all the eigenvectors and projected those same images onto the face space. For each image we then calculated the similarity to other images as follows. We used similarity measure from [43] and adapted it to [0,1] interval. Let a be the projection of image A onto ***W***, and let b be the projection of image B onto ***W***. Then similarity $\rho(A,B)$ of images A and B in feature space ***W*** can be expressed through their representations as:

$$\rho(A,B) = \frac{\langle a,b \rangle}{\|a\| \cdot \|b\|}, \tag{1}$$

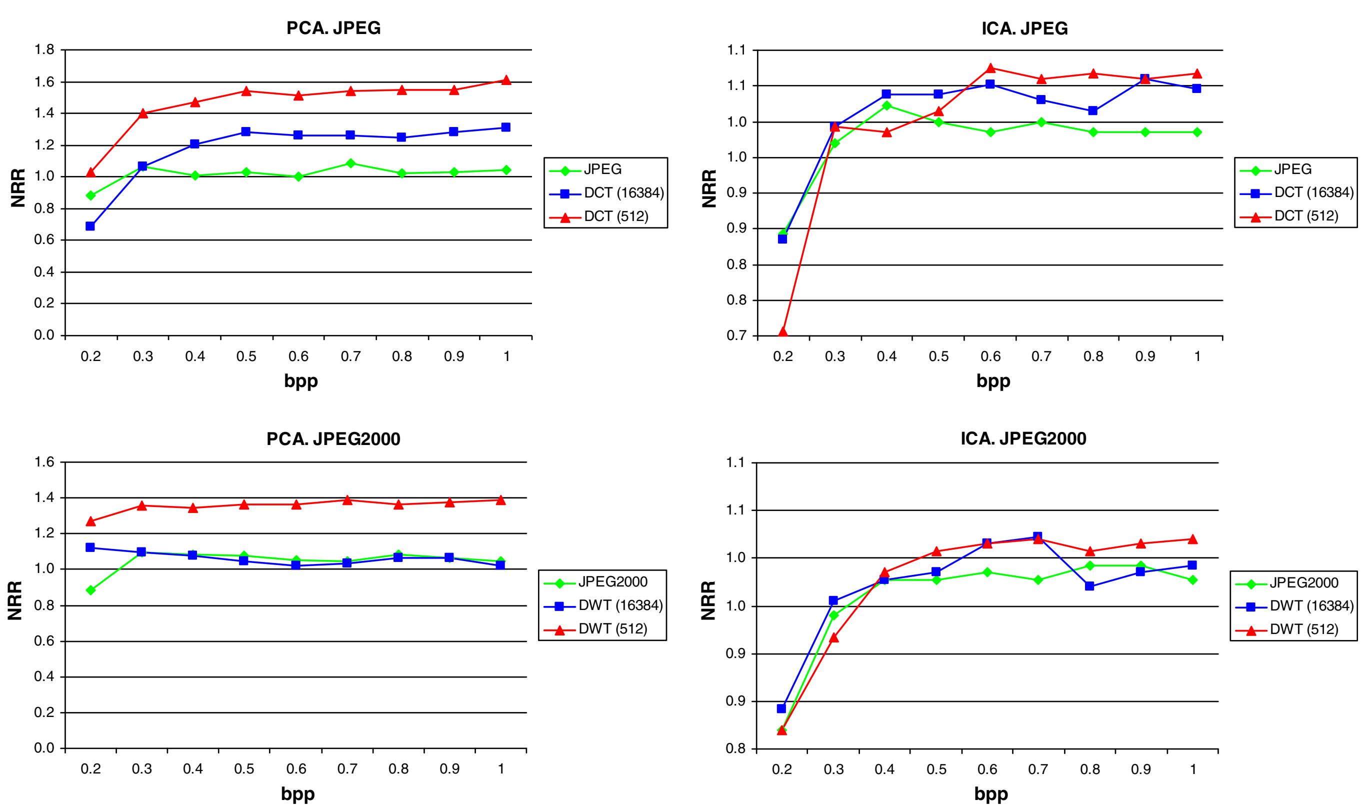

Fig. 4. NRR curves for the *fc* probe set for both compressed domain experiments.

where $\langle \cdot, \cdot \rangle$ is the inner product and $\| \cdot \|$ is the norm. Since the above formula is in essence the cosine of the angle between a and b it will yield values ranging from -1 to 1. We have adapted the formula to yield values in the range [0,1] so that 0 represents strong dissimilarity and 1 strong similarity (identical images):

$$\rho(A,B) = \frac{\langle a,b \rangle}{2 \cdot \|a\| \cdot \|b\|} + \frac{1}{2}. \quad (2)$$

The results of performing this experiment with uncompressed images (pixels) can be seen in Fig. 5(a), and performing the same experiment using DCT coefficients at 1 bpp in Fig. 5(b). Fig. 5 generally shows the similarity matrix where the diagonal terms have been removed for clearer display (they would always have the value of 1).

The lower the values for non-diagonal elements of the similarity matrix, the better the representation in feature space (regarding classification, i.e. recognition). In other words, since we used images of different subjects, we want the projections in feature space to be as dissimilar as possible. It is now obvious from Fig. 5 (a) and (b) that the projections of DCT coefficients are far less similar than the projections of pixels. This partly explains the frequent statistically significant improvements in RR for DCT coefficients seen in Table 1.

4.2.3. Computational complexity

We have shown in this section that working in compressed domain will not deteriorate performance and that compression can be used in face recognition systems. By using compression we have reduced the storage requirements and by doing recognition in compressed domain we have reduced the computational complexity (by avoiding the IDCT or IDWT). The computational complexity savings are twofold. In the training stage you can now have compressed images stored an only partly decompress them before training the algorithms. In the recognition stage, you can present only partly decompressed image to the system, thus again avoiding full decompression while retaining the advantages of storing and transmitting the compressed imaged instead of full uncompressed ones. In the following text we shall present a short numerical analysis of the savings achieved by working in compressed domain.

Let N be the total number of pixels in an image (or coefficients in the compressed domain). For our 128×128 images $N = 16384$. Computational complexity of IDCT is in the order of $O(N^2)$ in a standard implementation or $O(N \log N)$ when implemented through FFT. IDWT has computational complexity in the order of $O(N)$. Thus, avoiding the inverse transform in any scenario (training or recognition) saves up to $O(N \log N)$ computations in case of IDCT or $O(N)$ computations in case of IDWT.

All this draws one's attention to the possibilities of reducing the number of elements of the input vector that is presented to recognition algorithms as further means of possible savings in terms of the number of computations. To take PCA as an example for such savings (e.g. in the training stage), the approximate computational complexity of calculating the covariance matrix (computationally most intensive task in PCA process) is $O(N^3)$. Thus, for our 128×128 sized image, the computational complexity would be:

$$O(N^3) \approx O[(128 \times 128)^3] \approx O[(2^7 \times 2^7)^3] \approx O(2^{42}). \quad (3)$$

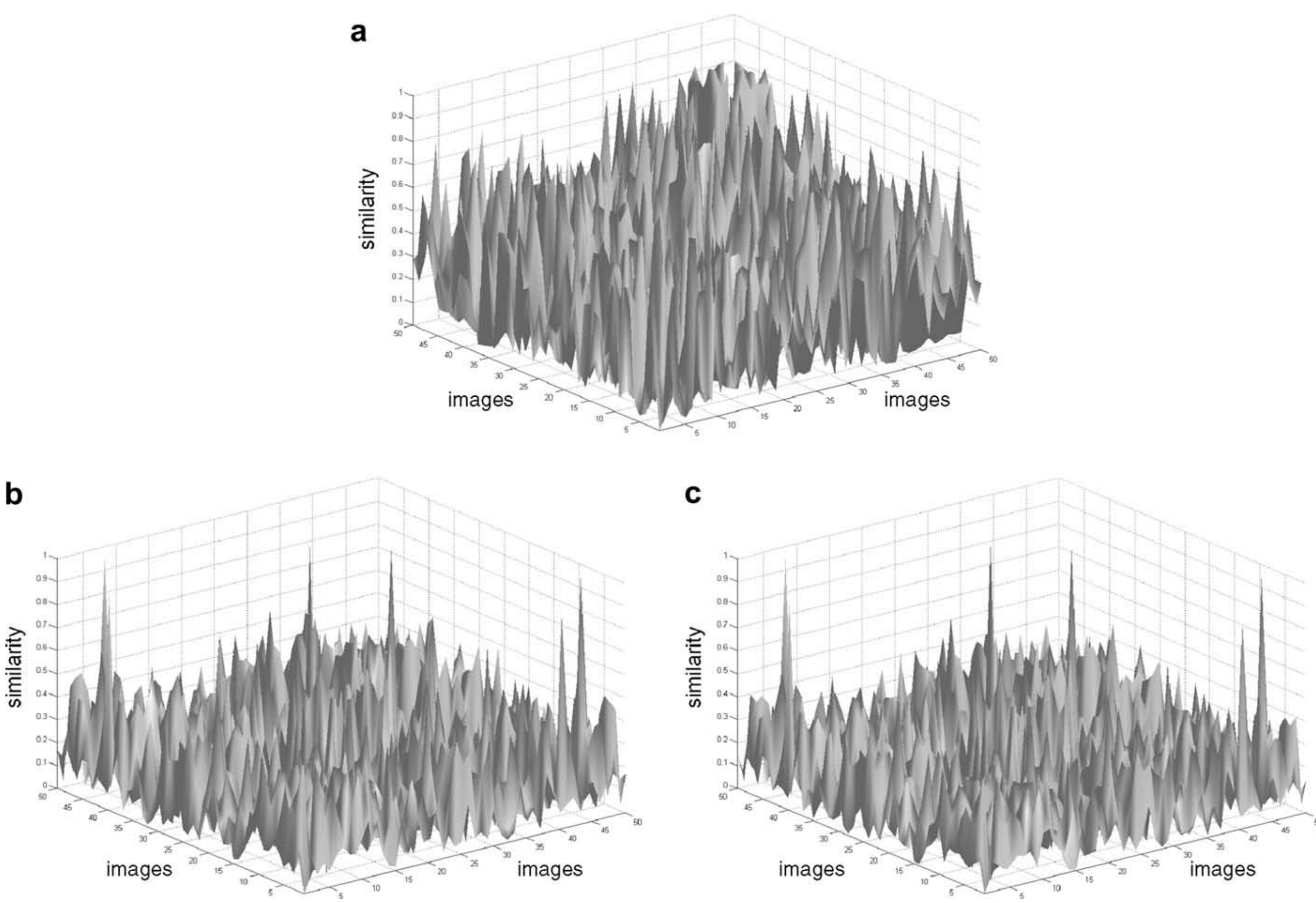

Fig. 5. Similarity matrix: (a) pixels; (b) all (16384) DCT coefficients; (c) preselected 512 coefficients.

Any reduction in the input vector that is normally sized $N \times 1$ is a major reduction in the number of necessary computations in the training stage.

If we were to move from the training stage to actual recognition, it can be shown that reducing the number of coefficients produces significant savings there as well. Continuing with PCA as an example, we can consider the stage where an unknown image (vector of coefficients in our case) is projected onto the feature space $\boldsymbol{W}$. Projection ω is calculated as:

$$\omega = W^T(I - \psi), \tag{4}$$

where I is the vector to be projected and ψ is the mean vector of all the training samples (please refer to [19] for details). For a vector of size $N \times 1$ and a 270-dimensional space (as in our experiments), according to (4) you should multiply two matrices of size: $[270 \times N] \times [N \times 1]$, where $N = 16384$. Finally, calculating any distance measure of a measure of similarity between two vectors is also proportional to their size, so reducing the size of vectors implies savings there as well.

In the following section we shall propose a simple yet effective method of coefficient preselection that will further reduce computational complexity of a hypothetical face recognition system operating in compressed domain.

5. Proposed coefficient preselection method

The proposed coefficient preselection method effectively captures the transform coefficients that are important for identification. This does not imply that those coefficients, when inversely transformed, would yield a reconstructed image that is similar to the uncompressed. Those coefficients also do not necessarily present the ones carrying the most energy or information.

Our basic hypothesis is that *coefficients varying the most for images of different faces are important for identifying subjects*. We used variance as a measure of how much a coefficient varies from image to image (i.e. from face to face) and considered the coefficients that are on the same spatial location in the images (see Fig. 6 for illustration of the proposed approach). To be able to measure the variance we had to make an alternative training set (T') to the one used in previous experiment. T' will only hold a subset of 225 images (one image per subject) that are taken under normal illumination conditions. This is important as we hope that such an approach would yield a subset of preselected coefficients that do not hold any information on the illumination changes. We argued that this could improve results for the *fc* probe set and, as will be shown in further results, it proved to be true. The new training set T' is now a subset of T ($T' \subset T$) with the number of images in T being $M = 225$. It will be used only to preselect the important intra-image coefficients. In the testing stage the same training set T will be used, as in the previous section, in order to be able to directly compare the results.

All the images in T' are compressed and then partly decompressed to access the transform coefficients (as in the previous experiment, illustrated in Fig. 2). The variance of each individual coefficient is calculated as follows: let (u,v) be the spatial coordinates of a coefficient and let a single coefficient extracted from the decompression procedure be $\widehat{F}(u,v)$; the variance $\sigma^2_{u,v}$ for $\widehat{F}(u,v)$ is then calculated as:

$$\sigma^2_{u,v} = \frac{1}{M-1}\sum_{i=1}^{M}[\widehat{F}(u,v) - \overline{F}(u,v)]^2, \tag{5}$$

where $\overline{F}(u,v)$ is the mean value of $\widehat{F}(u,v)$ for all images in T'. Using (5) we calculated the variance for all N coefficients (16384) and kept the 512 coefficients with the largest variance. We have chosen to keep 512 coefficients as a rule of thumb and a compromise between reducing the N, and keeping the number of coefficients as an order of 2 (for faster computation). Additionaly, results of numerous experiments we performed with other values (not reported here) give lower RR on average. The same procedure was applied for DCT and DWT coefficients.

The vector that is the input to recognition algorithms (both in the training and in the testing stage) is now $N \times 1 = 512 \times 1$ instead of 16384×1. The spatial location of the preselected coefficients is shown in Fig. 7. The location of the preselected DCT coefficients is shown on the left side of Fig. 7, and the location of the preselected DWT coefficients is shown on the right. The coefficients with the highest intra-image variance are roughly located around the face features (eyes, nose and mouth), both in case of DCT and DWT. Surprisingly, in case of DWT, the important coefficients seem to be mainly in the diagonal (HH) band that captures the least energy.

Using the preselected 512 coefficients as input to recognition algorithms, the computational complexity formulated in (3) now becomes:

$$O(N^3) \approx O[(512)^3] \approx O[(2^9)^3] \approx O(2^{27}). \tag{6}$$

This shows a significant reduction in computational complexity in the training stage. At recognition stage, the matrices to be multiplied are now according to (4) the size of $[270 \times N] \times [N \times 1]$, where $N = 512$.

5.1. Experimental setup

To be able to objectively compare results from this experiment to the results from the previous section, we shall use exactly the same experimental setup (Fig. 3) along with the same T, G and P sets. The only real difference is that we now use the preselected 512 coefficients instead of all 16384. The alternative training set

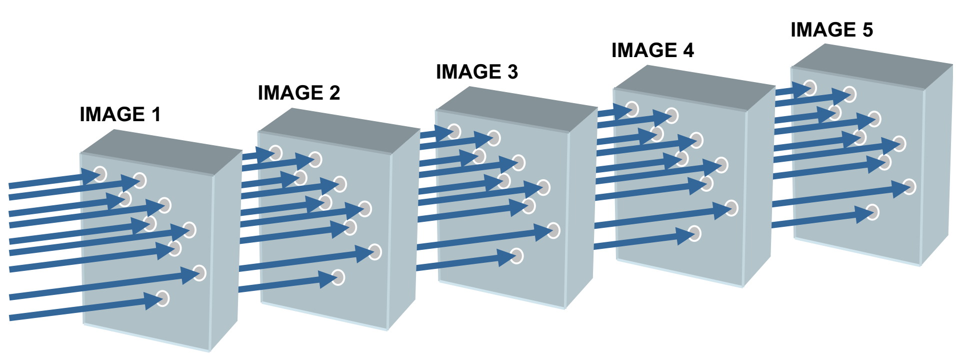

Fig. 6. Illustration of the proposed coefficient preselection method. The square blocks represent images and the dots represent coefficients. The same spatially located coefficients are analyzed across all images in the training set.

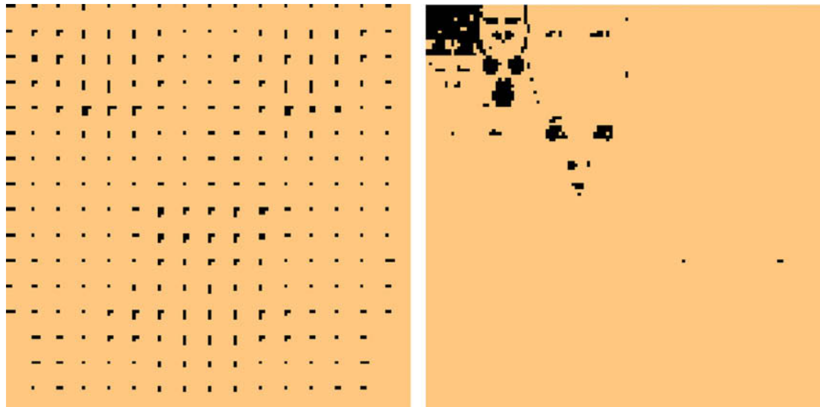

Fig. 7. Spatial location of the preselected 512 coefficients within the image compressed to 1 bpp; left - the JPEG compression scheme (DCT coefficients); right - the JPEG2000 compression scheme (DWT coefficients).

T' is only used to determine the spatial location of the coefficients with high variance, and then the standard training set T is used for the actual PCA and ICA training.

5.2. Results and analysis

The results of this experiment can be seen in Tables 3 and 4 and Fig. 4 (DCT(512) and DWT(512)). Again, as in the previous experiment, we can generally conclude that *using only preselected 512 transform coefficients per image does not significantly deteriorate performance*. We shall in the following text give a detailed analysis of the results.

5.2.1. Analysis by probe sets

For the *fb* probe set and DCT coefficients (Table 3) we can see a significant increase in RR for PCA compared both to RR with uncompressed images and to results of the previous experiment. At 1 bpp there is a 2.2% increase in RR compared to previous experiment and a significant difference of 3.1% compared to uncompressed images. Taking other compression rates into account, we can conclude that coefficients preselection has without a doubt improved results in this case. ICA also exhibits a significant RR increase for this probe set and DCT coefficients. This is most obvious at 1 bpp where the RR difference to uncompressed images is now 1.4% and is statistically significant according to McNemar's test. For the DWT coefficients (Table 4) PCA now has RR significantly higher at all compression rates. The worst result in the previous experiment (77.8% at 1 bpp - significantly worse than with uncompressed images) is now increased to 81.7% and this increase is statistically significant. ICA yielded results quite similar to previous experiment, with one exception at 0.3 bpp with the increase in RR from 83.8% to 84.4%. It is interesting to observe that both PCA and ICA perform best at 0.3 bpp in this experiment with 512 preselected DWT coefficients.

Results for the *fc* probe set and DCT coefficients can be seen in Table 3. RR for PCA is significantly increased compared to uncompressed images and to RR in the previous experiment. The increase is not only statistically significant but enormous in absolute value as well (29.4% higher than with uncompressed images and 14.4% higher than with all coefficients at 1 bpp). This is also the only case where ICA is not a superior method in total RR (ICA has the highest RR of 73.2% at 1 bpp while PCA has the highest RR of 77.3% at 1 bpp). As can be seen in Table 3 and in NNR curves in the top left graph in Fig. 4, the results for PCA drop towards higher compression rates and RR is at 0.2 bpp only slightly (and insignificantly) higher than with uncompressed images. ICA also shows a slight (at 1 bpp even statistically significant) increase in RR, but drops faster than PCA towards higher compression rates (top right graph in Fig. 4). The effect of significant increase in RR for PCA can also be seen in the JPEG2000 compressed domain in Table 4 (using the DWT coefficients). Although, the absolute difference in RR is now only 18.6% compared to uncompressed images and 17.5% compared to all coefficients. The RR with DWT shows more stability as it now remains significantly higher even at 0.2 bpp. This is also evident from the NRR curves (bottom left graph in Fig. 4). ICA now shows a slight improvement in RR only at 0.5 bpp, whereas the RR is slightly lower at other compression rates. However, by examining the bottom right graph in Fig. 4, we can see that ICA drops significantly only below 0.3 bpp. The overall increase in RR both for DCT and DWT coefficients and for both recognition methods in this experiment should not come as a surprise. It is exactly what we anticipated and hoped to achieve by our proposed coefficients preselection method where we only considered images taken under normal illumination. We hinted in the part explaining the proposed method that the coefficients selected this way will not carry the information on illumination and that this should increase the RR for the *fc* probe set. This presumption is now confirmed by this experiment and it proves the appropriateness of the proposed approach, especially the part on what the T' should look like.

For *dup1* and *dup2* probe sets with DCT coefficients, PCA shows a slight decrease in RR but at most of the compression rates the difference is not statistically significant. At 1 bpp the RR is for both probe sets higher than with uncompressed images. For *dup1* probe set ICA performs slightly worse than with uncompressed images, but still the difference statistically insignificant. Significant increase in RR is observed at 1 bpp for the *dup2* probe set, while the RR at other compression rates remains slightly above the one obtained with uncompressed images and stable. For the DWT coefficients PCA results are similar to the ones obtained using all coefficients, except that now the differences are insignificant even at 0.2 bpp. RR for ICA with DWT increased significantly at 1 bpp and 0.5 bpp, giving the absolute difference in RR of 5.1%. As in the previous experiment (section 4), the results seem more stable when using DWT than DCT for these two probe sets.

5.2.2. Analysis in feature space

Result of performing the same experiment as in section 4.2.2 with the preselected 512 coefficients can be seen in Fig. 5(c). The similarity between projections of different subjects is again smaller than with the uncompressed images (pixels; Fig. 5(a) and quite similar to Fig. 5(b), where all the coefficients are used. This is an important observation as it proves that there is little to none disturbance in feature space, even if the vectors now have only 512 elements. In other words, *we have gained significant computational* *performance savings while retaining practically the same relationships between projections in feature space.*

5.2.3. Computational time savings

As we have analyzed the numerical benefits associated to computational complexity of the proposed approach in the first part of section 5, we shall here give only a small practical analysis of the computations of the last experiment. The computations necessary for the training of PCA method took about 5.2 s when using either coefficients or uncompressed images (as the number of elements and vector sizes are the same in both cases), and less than 1 s for the proposed method (under the same conditions, the same variables preloaded in memory etc., on a Pentium 4 at 2.4 GHz with 1 GB RAM, Matlab 7.0 R14).

5.3. Compressed domain face recognition system

A hypothetical face recognition system working in compressed domain is shown in Figs. 8 and 9. Analyzing the results of the experiments we have concluded that ICA is the overall more accurate than PCA and this is why we have chosen it as a recognition method in this hypothetical system.

Fig. 8. represents a system in the training stage, with two important presumptions: firstly, the images to be used in this stage are uncompressed, and secondly, the face detection in compressed domain is considered to be solved and is omitted from the analysis. Face detection and normalization in compressed domain is possible and an example can be seen in [44].

The original images from the training set of size $m \times n$ are normalized to $m' \times n'$ in the "Norm." module and then compressed ("Comp.") and partly decompressed ("Decomp."). The subset T' is then used as input to the "VAR." module which outputs the spatial location of the 512 coefficients to be preselected in the "Selection" module. Training set T is then used (after the selection process) to perform PCA (deriving subspace $\boldsymbol{W}_{\mathrm{PCA}}$ which is used as a preprocessing step to ICA. ICA will consequently output the subspace $\boldsymbol{W}_{\mathrm{ICA}}$ in which the actual recognition will be done through distance calculation. Once the subspace is calculated, the gallery images (actually the preselected coefficients of the gallery images) are projected onto it and the projections $(\Omega_1, \ldots, \Omega_G)$ are stored in a database. If the images in the training set or the gallery set are already compressed, the "Comp." module can simply be omitted from Fig. 8.

Fig. 9 illustrates the recognition stage in which an already compressed probe image of size $m'' \times n''$ is normalized in compressed domain (yielding an $m' \times n'$ image) and then partly decompressed. The coefficients are preselected using the information on the location of the 512 important ones and they are projected onto the $\boldsymbol{W}_{\mathrm{ICA}}$ subspace. Distance between this projection and all the stored ones is calculated and the minimal distance (using the cosine angle metric), i.e. the identity of the closest gallery match, is what the system outputs as the identity of the unknown probe image.

6. Conclusion

This paper presents the first systematic and detailed research of the possibilities of implementing face recognition methods directly in JPEG and JPEG2000 compressed domain. We have examined all aspects that can arise in such a scenario, ranging from the effects of compression on recognition rate to technical implementation issues like computational time and storage requirements savings. We have performed numerous experiments using two standard recognition methods (PCA and ICA) and statistically proved that working directly in compressed domain does not deteriorate recognition results significantly. All the results from the experiments in compressed domain described in this paper were compared to the results in pixel domain (using either uncompressed or fully decompressed images).

We have described and proposed two methods in compressed domain: one where all the available transform coefficients were used, and one where we preselected only a part of coefficients (the ones with the largest intra-image variance for different subjects). Our experimental results on the FERET database for both methods show that the differences in recognition rate in compressed domain compared to uncompressed images (pixel domain) are roughly around a couple of percentages. The recognition rates are in many cases higher in compressed domain and we have proved that this increase is significant by using McNemar's hypothesis test. Our proposed method of preselecting only a part of transform coefficients in compressed domain proved to be more accurate than just using all the available coefficients. The improvements in recognition rates compared to pixel domain in this case were more frequent and statistically more significant.

The results presented in this paper prove that efficient face recognition in compressed domain is possible and that the real-life technical implementation of such a system would result in considerable computational time and storage requirements savings.

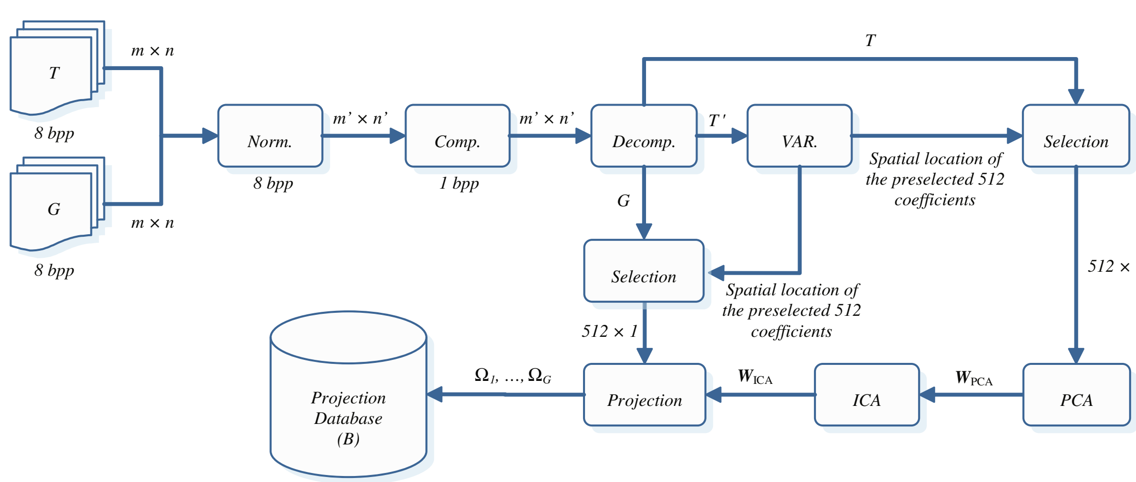

Fig. 8. A hypothetical face recognition system working directly in compressed domain - *training stage*; compression at 1 bpp; ICA used as recognition method.

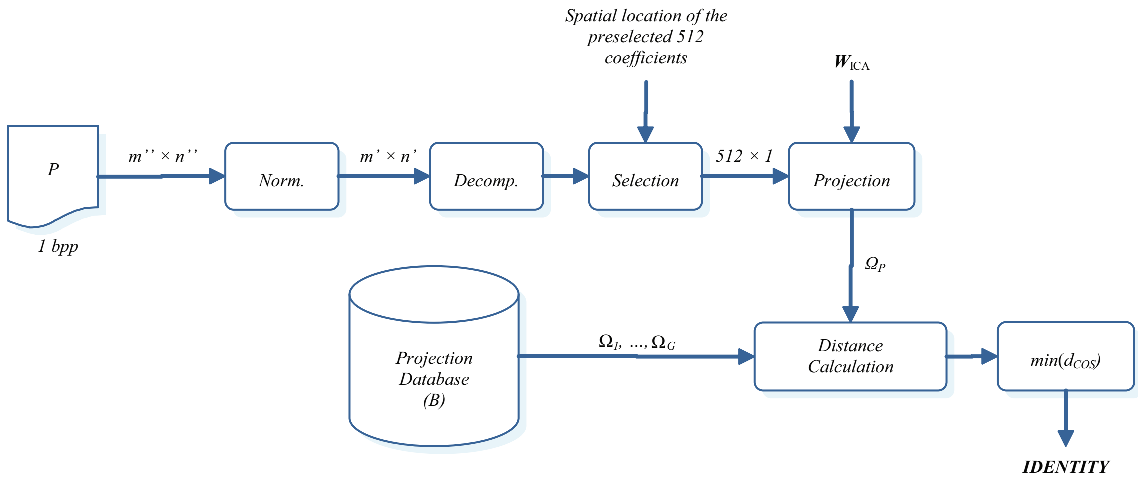

Fig. 9. A hypothetical face recognition system working directly in compressed domain - *recognition stage*; compression at 1 bpp; ICA used as recognition method.

Acknowledgement

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 paper 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, A. Rosenfeld, P.J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys 35 (4) (2003) 399–458.
[2] K. Delac, M. Grgic, A Survey of Biometric Recognition Methods, Proc. of the 46th International Symposium Electronics in Marine, ELMAR-2004, Zadar, Croatia, 16-18 June 2004, pp. 184-193.
[3] S.Z. Li, A.K. Jain (Eds.), Handbook of Face Recognition, Springer, New York, USA, 2005.
[4] K. Delac, M. Grgic (Eds.), Face Recognition, I-Tech Education and Publishing, Vienna, 2007.
[5] S. Rakshit, D.M. Monro, An Evaluation of Image Sampling and Compression for Human Iris Recognition, IEEE Trans. on Information Forensics and Security 2 (3) (2007) 605–612.
[6] S. Matschitsch, M. Tschinder, A. Uhl, Comparison of Compression Algorithms' Impact on Iris Recognition Accuracy, Lecture Notes in Computer Science 4642 (2007) 232.
[7] W. Funk, M. Arnold, C. Busch, A. Munde, Evaluation of Image Compression Algorithms for Fingerprint and Face Recognition Systems, Proceedings from the Sixth Annual IEEE Systems, Man and Cybernetics (SMC) Information Assurance Workshop, 2005, pp. 72-78.
[8] A. Mascher-Kampfer, H. Stogner, A. Uhl, Comparison of Compression Algorithms' Impact on Fingerprint and Face Recognition Accuracy, Proceedings of SPIE, Vol. 6508, Visual Communications and Image Processing, 2007, 12 pages.
[9] M. Elad, R. Goldenberg, R. Kimmel, Low Bit-Rate Compression of Facial Images, IEEE Trans. on Image Processing 16 (9) (2007) 2379–2383.
[10] Face Recognition Format for Data Interchange, ANSI INCITS 385-2004, American National Standard for Information Technology, New York, 2004.
[11] Biometric Data Interchange Formats - Part 5: Face Image Data, ISO/IEC JTC1/SC37 N506, ISO/IEC IS 19794-5, 2004.
[12] D.P. McGarry, C.M. Arndt, S.A. McCabe, D.P. D'Amato, Effects of Compression and Individual Variability on Face Recognition Performance, Proc. of SPIE 5404 (2004) 362–372.
[13] M.K. Mandal, F. Idris, S. Panchanathan, A Critical Evaluation of Image and Video Indexing Techniques in the Compressed Domain, Image and Vision Computing 17 (7) (1999) 513–529.
[14] G.K. Wallace, The JPEG Still Picture Compression Standard, Communications of the ACM 34 (4) (1991) 30–44.
[15] A. Skodras, C. Christopoulos, T. Ebrahimi, The JPEG 2000 Still Image Compression Standard, IEEE Signal Processing Magazine 18 (5) (2001) 36–58.
[16] D.M. Blackburn, J.M. Bone, P.J. Phillips, FRVT 2000 Evaluation Report, 2001, available at: www.frvt.org/FRVT2000/documents.htm.
[17] K. Delac, M. Grgic, S. Grgic, Effects of JPEG and JPEG2000 Compression on Face Recognition, Lecture Notes in Computer Science - Pattern Recognition and Image Analysis 3687 (2005) 136–145.
[18] K. Delac, M. Grgic, S. Grgic, Image Compression Effects in Face Recognition Systems, in: K. Delac, M. Grgic (Eds.), Face Recognition, I-Tech Education and Publishing, Vienna, 2007, pp. 75–92.
[19] M. Turk, A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neurosicence 3 (1) (1991) 71–86.
[20] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face Recognition by Independent Component Analysis, IEEE Trans. on Neural Networks 13 (6) (2002) 1450–1464.
[21] P.J. Phillips, H. Wechsler, J. Huang, P.J. Rauss, The FERET Database and Evaluation Procedure for Face Recognition Algorithms, Image and Vision Computing 16 (5) (1998) 295–306.
[22] 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 22 (10) (2000) 1090–1104.
[23] M. Shneier, M. Abdel-Mottaleb, Exploiting the JPEG Compression Scheme for Image Retrieval, IEEE Trans. on Pattern Analysis and Machine Intelligence 18 (8) (1996) 849–853.
[24] W.B. Seales, C.J. Yuan, W. Hu, M.D. Cutts, Object Recognition in Compressed Imagery, Image and Vision Computing 16 (5) (1998) 337–352.
[25] S. Eickeler, S. Muller, G. Rigoll, High Quality Face Recognition in JPEG Compressed Images, Proc. of the 1999 International Conference on Image Processing, ICIP'99, Vol. 1., Kobe, Japan, 24-28 October 1999, pp. 672-676.
[26] S. Eickeler, S. Muller, G. Rigoll, Recognition of JPEG Compressed Face Images Based on Statistical Methods, Image and Vision Computing 18 (4) (2000) 279–287.
[27] Z.M. Hafed, M.D. Levine, Face Recognition Using the Discrete Cosine Transform, International Journal of Computer Vision 43 (3) (2001) 167–188.
[28] R. Tjahyadi, W. Liu, S. Venkatesh, Application of the DCT Energy Histogram for Face Recognition, Proc. of the 2nd International Conference on Information Technology for Applications, ICITA 2004 (2004) 305-310.
[29] W. Chen, M.J. Er, S. Wu, PCA and LDA in DCT Domain, Pattern Recognition Letters 26 (15) (2005) 2474–2482.
[30] C.L. Sabharwal, W. Curtis, Human Face Recognition in the Wavelet Compressed Domain, Smart Engineering Systems, ANNIE 97, St. Louis, Missouri, USA, 7 (1997) 555-560.
[31] C. Garcia, G. Zikos, G. Tziritas, Wavelet Packet Analysis for Face Recognition, Image and Vision Computing 18 (4) (2000) 289–297.
[32] J.T. Chien, C.C. Wu, Discriminant Waveletfaces and Nearest Feature Classifiers for Face Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence 24 (12) (2002) 1644–1649.
[33] B. Li, Y. Liu, When Eigenfaces are Combined with Wavelets, Knowledge-Based Systems 15 (5) (2002) 343–347.
[34] B.L. Zhang, H. Zhang, S.S. Ge, Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory, IEEE Trans. on Neural Networks 15 (1) (2004) 166–177.
[35] H.K. Ekenel, B. Sankur, Multiresolution Face Recognition, Image and Vision Computing 23 (5) (2005) 469–477.
[36] K. Delac, M. Grgic, S. Grgic, Towards Face Recognition in JPEG2000 Compressed Domain, Proceedings of the 14th International Workshop on Systems, Signals and Image Processing (IWSSIP) and 6th EURASIP Conference focused on Speech & Image Processing, Multimedia Communications and Services (EC-SIPMCS), Maribor, Slovenia, 27-30 June 2007, pp. 155-159.
[37] K. Delac, M. Grgic, S. Grgic, Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set, International Journal of Imaging Systems and Technology 15 (5) (2006) 252–260.

[38] P. Fonseca, J. Nesvadha, Face Detection in the Compressed Domain, ICIP 2004 3 (24-27) (2004) 2015–2018.
[39] Official site of the Joint Photographic Experts Group (JPEG) and Joint Bi-level Image experts Group (JBIG), Available: www.jpeg.org.
[40] Q. McNemar, Note on the Sampling Error of the Difference Between Correlated Proportions or Percentages, Psychometrika 12 (2) (1947) 153–157.
[41] J.R. Beveridge, K. She, B.A. Draper, G.H. Givens, A Nonparametric Statistical Comparison of Principal Component and Linear Discriminant Subspaces for Face Recognition, Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR'01, Kaui, HI, USA, December 2001, pp. 535-542.
[42] J.R. Beveridge, K. She, B. Draper, G.H. Givens, Parametric and Nonparametric Methods for the Statistical Evaluation of HumanID Algorithms, IEEE Third Workshop on Empirical Evaluation Methods in Computer Vision, Kauai, HI, December 2001.
[43] G.C. Feng, P.C. Yuen, D.Q. Dai, Human Face Recognition using PCA on Wavelet Subband, Journal of Electronic Imaging 9 (2) (2000) 226–233.
[44] H. Luo, A. Eleftheriadis, On Face Detection in the Compressed Domain, Proc. of the 8th ACM International Conference on Multimedia, Marina del Rey, CA, USA, 30 October - 03 November 2000, pp. 285-294.