Paweł FORCZMAŃSKI[1], Georgy KUKHAREV[1], Nadezhda L'vovna SHCHEGOLEVA[2]

West Pomeranian University of Technology, Szczecin, Faculty of Computer Science and Information Systems (1),
Saint Petersburg State Electrotechnical University (LETI) (2)

An algorithm of face recognition under difficult lighting conditions

Abstract. *The paper addresses the problem of face recognition for images with lighting problems – flashes, shadows and very low brightness level. Presented algorithm, allowing to eliminate above problems, is based on 2DDCT (two-dimensional Discrete Cosine Transform) supported by brightness gradient reduction, reduction of spatial low frequency spectral components and fusion of spectral features conditioned on average intensities. Presented experiments were conducted on image databases Yale B and Yale B+.*

Streszczenie*. W artykule zaprezentowano zadanie rozpoznawania twarzy na podstawie obrazów uzyskanych w trudnych warunkach oświetleniowych, posiadających odbłyski, cienie i niski poziom jasności. Zaprezentowany algorytm pozwala na wyeliminowanie wymienionych problemów i bazuje na dwuwymiarowej dyskretnej transformacie kosinusowej połączonej z redukcją gradientu jasności, eliminację niskoczęstotliwościowych komponentów widma i fuzji komponentów widma zależnej od średniej jasności obrazu Jako uzupełnienie, przedstawiono eksperymenty przeprowadzone na bazach Yale B i Yale B+. (****Algorytm rozpoznawania twarzy w trudnych warunkach oświetleniowych****)*

Keywords: face recognition, illumination reduction, two-dimensional Discrete Cosine Transform
Słowa kluczowe: rozpoznawanie twarzy, eliminacja oświetlenia, dwuwymiarowa dyskretna transformata kosinusowa

Introduction

One of the most crucial problems in face recognition practice is the variations of light intensity in input images processed by FaReS (Face Recognition System). In such situation we have to deal with two types of complications in the area of face and its background. The first one is related to local shadows, while the second one is associated with so called global shadows. Local shadows change the form of individual parts of face (nose, mouth, and eyes) and distort the boundaries of face area. Global shadows significantly reduce the discrimination of various face areas against general background and/or completely hide them.

Such kinds of problems strongly influence the accuracy of FaReS operation, thus this is the main reason of lasting interest of face recognition specialists [1-11]. Analysis of the literature leads to the observation, that the problem is still unsolved in a satisfactory way. There is only a few methods oriented at this problem, namely:

1. Processing the images in order to equalize brightness variation (intensity equalization); reduction of intensity gradient (gamma correction, logarithmic transformation), invariant (in respect to intensity) image representation using the LBP - Local Binary Patterns, and LTV – Logarithmic Total Variation);
2. Representation of face images with lighting problems (FILP) using the eigenbase decompositions and corresponding models based on Eigenface approach;
3. Representation of FILP with spectral features using the wavelets and cosine transformation with elimination of low-frequency components;
4. Extension of FaReS database with new patterns having all distortions related to lighting problem of face images;

In real-world situations we have a limited training database, thus we should consider the task of FILP recognition only in the case when FaReS base is not extended with patterns having all possible variants of global and/or local shadows [4, 6, 9] in contrast to setting defined in [1-3, 5, 7, 8].

One of the most promising, yet computationally expensive approaches related to eigenbasis representation employs Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) [1-3], and also Canonical Correlation Analysis (CCA) [6], together with Discrete Cosine Transform (DCT) as FILP preprocessing step. There are also several approaches to the problem of illumination compensation involving wavelets, e.g. [10] or very complex models related to pose estimation and further illumination handling [11]. In this paper we focus on methods involving dimensionality reduction approach. The authors of [1] show that in order to solve recognition task using the PCA and LDA, FILPs should be transformed into spectral features using two-dimensional DCT (2DDCT). At the same processing stage, the low frequency spectral components are removed, as corresponding to „shadow" components. In [1] the authors proposed the following procedure involving one-dimensional PCA and LDA:

(1) $$Log(X) \rightarrow 2DDCT \rightarrow C \rightarrow C\text{-}C_{low} \rightarrow 1DPCA;$$
$$Log(X) \rightarrow 2DDCT \rightarrow C \rightarrow C\text{-}C_{low} \rightarrow 1DLDA,$$

where: X – FILP image, $Log(X)$ – logarithm of pixel intensities; C – spectrum in the basis of two-dimensional DCT (2DDCT); $C\text{-}C_{low}$ – discarding low frequencies.

In [2] the authors presented another procedure, where the spectral representation of FILP is transformed back into original intensity representation (inverse 2DDCT) just before PCA step (which, according to authors' claims, increases the recognition rate of FILP):

(2) $$Log(X) \rightarrow 2DDCT \rightarrow C \rightarrow C\text{-}C_{low} \rightarrow Inv2DDCT \rightarrow 1DPCA;$$

Similar approach was also presented in [5], however, calculation related to PCA is optimized by the Gramm matrix evaluation.

This paper presents an algorithm of solving lighting problems in the aspect of facial portraits recognition, which is much more simple in comparison to the above presented approaches, yet its efficiency is similar. It uses the following methods:

- The template database of FaReS does not include images with local and/or global shadows;
- It uses simple preprocessing of original images aimed at reduction of intensity gradient (gamma correction and intensity logarithm);
- It uses 2DDCT as the only instrument of transformation of original data into the space of spectral features;
- Spectral features are combined at classification step, in case of low intensity level.

Characteristics of images used in experiments

The literature review shows that most of the reported experiments in the area of facial portrait recognition in the presence of variable lighting conditions are conducted on Yale database [12] as it seems to be a de facto standard in the scientific community. Complete Yale database includes original Yale B images and its extension called Yale B+ [12]. In experiments we used 2452 out of 2470 images from Yale B and Yale B+ sets, containing the central part of face area of 38 subjects (18 images were omitted since they cannot be read from files published on web site [12]). All images are stored in grayscale in matrices of 192×168 pixels, divided into 6 sets, labeled Subset 0 ÷ Subset 5, respectively. Images in Subset 0 have no blinks, shadows and feature ambient lighting. Images in Subset 1 ÷ Subset 5 were obtained by modeling the spatial movement of a light source, hence contain various variants of shadows – flashes (Subset 1 and Subset 2), local shadows (Subset 3) and lateral shadows (Subset 4 and Subset 5), as well as global shadows (Subset 5). The most difficult for recognition are images from Subset 3 ÷ Subset 5. Figure 1 presents example images from these sets.


Fig.1. Selected images from Yale database

Face image preprocessing
It is obvious that without brightness equalization of test images the recognition rate will be very low. As showed in [1-9], the methods of brightness enhancing (shown in Fig. 2) can employ gamma correction or performing a logarithm on pixel intensities. However, this procedure must respect such parameters of input image as mean value, local brightness, boundaries of shadows, as well as contrast. Unfortunately, such procedure may have a negative influence on the recognition accuracy. Figure 2 presents original images influenced by different lighting conditions together with results of applying one of two enhancing procedures: gamma correction (G) and brightness logarithm (Log). Observed distortions (like not removed shadows, introduced new bright spots, loss of contrast, noise) are clearly visible in resulted images – especially when compared to original images. Even though, the boundaries of different parts of face can be easily detected, and the anthropometric parameters of faces may be successfully explored. It is also visible that gamma correction and logarithmic transformation of original image unveil different parts of face, originally hidden in the shadow, leading to the improvement of recognition rate.

Both procedures can be easily described using the following formulas. Let I be an image of size $M\times N$ pixels, containing values from the range $<0,\ 255>$. The gamma correction procedure applied for changing the brightness of each pixel i of image I is implemented as follows:

$$i_G(m,n) = i(m,n)^{\gamma},\ \forall\ m=1,\ldots,M \text{ and } n=1,\ldots,N, \tag{3}$$

where: $i_G(m,n)$ – pixel after correction; γ - coefficient of power transform, $\gamma << 1$.

Logarithmic transformation consists of two steps. In the first step all zero values in the image matrix are replaced by ones, so that:

$$i(m,n) = i(m,n)+1,\ \forall\ i(m,n) \equiv 0. \tag{4}$$

In the next step the logarithm is calculated:

$$i_L(m,n) = \log(i(m,n)),\ \forall\ m=1,\ldots,M \text{ and } n=1,\ldots,N, \tag{5}$$

where $i_L(m,n)$ – a new brightness of a pixel.

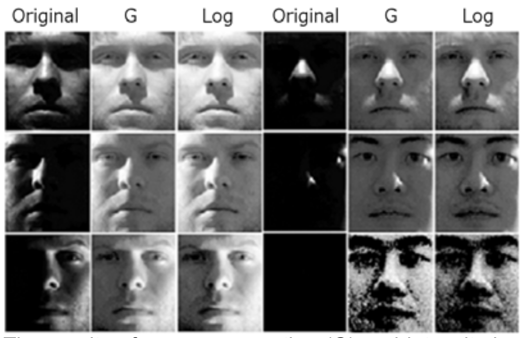

Fig.2. The results of gamma correction (G) and intensity logarithm (Log) applied on original images

In many previous works further stages of preprocessing are based on PCA/KLT approach. Here we employ much simpler approach, namely DCT, since the basis functions of cosine transform are optimal (in the sense of minimal loss of energy) for approximation of eigenfunctions for the whole population (or just large collections) of digital images. As the cosine transformation represents original face images with relatively small number of features, we exclude the entire preprocessing step based on PCA/KLT. For an adequately precise reconstruction of the face (in case of an image with uniform lighting) it is enough to take not more than 20 spectral components, selected from upper left corner of spectrum matrix [13].

Practical implementation of 2D DCT
We can define the 2DDCT in a matrix form as follows [13]:

$$Y = F_1\, I\, F_2, \tag{6}$$

where I – original image of size $M\times N$; Y – transformation result; F_1 and F_2 – projection matrices of size $d_1\times M$ and $N\times d_2$, where:

$$f(d_1,m) = \begin{cases} 1/\sqrt{M}, & \text{if } d_1 = 0, \\ \sqrt{\dfrac{2}{M}}\cos\dfrac{\pi(2m+1)d_1}{2M}, & \text{if } d_1 = 1,\ldots,D_1-1, \end{cases} \quad \forall\, m = 0,1,\ldots,M-1. \tag{7}$$

$$f(n,d_2) = \begin{cases} 1/\sqrt{N}, & \text{if } d_2 = 0, \\ \sqrt{\dfrac{2}{N}}\cos\dfrac{\pi(2n+1)d_2}{2N}, & \text{if } d_2 = 1,\ldots,D_2, \end{cases} \quad \forall\, n = 0,1,\ldots,N-1. \tag{8}$$

One important spectrum's feature should be noted. Two-dimensional DCT allows to obtain spatial-spectral features, invariant against mirror – symmetric modification of a head to the left and/or right in the original image plane. This fact is proved by a result presented in Fig. 3. It shows two images from Yale B set, featuring mirror symmetry of lighting source. Original spectra correspond to regular (signed) 2DDCT of both images (one is represented by dots while the other with solid line). Absolute values of 2DDCT spectra for these images are identical. Thus, the absolute value of 2DDCT spectrum is invariant against mirror-symmetric transformation of face images. Please note that in practical applications of face recognition above mentioned fact of invariance can also result in better face recognition in case of head rotation around vertical axis, and also in case of not perfect mirror-symmetry of lighting source.

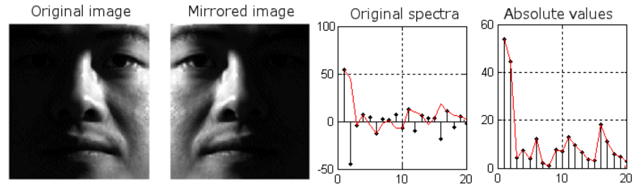

Fig.3. Sample face image and its mirrored version and the comparison of their spectra (regular and absolute values)

Processing algorithm

Developed algorithm of processing and recognition of FILP consists of five following stages, as presented in Fig. 4:

1. Construction of two projection matrices F_1 and F_2 of sizes $d_1 \times M$ and $N \times d_2$, respectively (here $d_1 << M$ and $d_2 << N$; $d_1 \neq d_2$ in general case).
2. Application of 2DDCT for templates with a preliminary processing, according to (3)-(5), resulting a database of processed templates. Note, that the projection of original data – images of size $M \times N$ – into new feature space, resulted in a new representation with spectral matrices Y of size $d_1 \times d_2 << MN$, so that: $Y = F_1 \tilde{I} F_2$, where $\tilde{I}$ – image after preprocessing stage.
3. Preprocessing of test image by means of gamma correction (3) or gamma correction (3) followed by logarithmic correction (4) and (5). The choice of preprocessing type depends on average brightness J of test image. If value of J is lower than a certain threshold P, then both optional steps of preprocessing are executed, otherwise only one step – i.e. gamma correction (3) or logarithmic correction (4) and (5) is employed.
4. Projection of transformation result of previous step into a new feature space based on 2DDCT. Projection result consists of spectral matrices (precisely, one or two matrices) of size $d1 \times d2 << MN$. Final projection result (steps 2 and 4) is presented as a vector, read from top left corner of spectral matrices. Number of elements of such vector is equal to $d \times (d+1)/2$, where $d = min(d1, d2)$.
5. Classification of the result obtained in step 4 through comparison with features of images from all known classes, stored in the training set (results of step 2). A classification criterion is the minimum distance in metric $L1$ ($MDC/L1$). Test image is assigned to a class (one of known classes) having a minimum distance ($rank=1$).

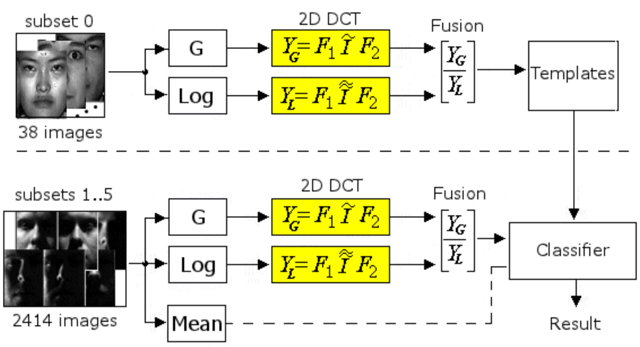

Fig.4. Structure of FILP recognition system

In the classification process, the brightness J of test image is evaluated. If $J<P$, then the projection results for such image obtained in step 4 of the algorithm are fused in a one common vector, which is compared to respective collection of features obtained for all templates. In our experiments, a value of P was set to 50, as corresponding to the cases with shadows covering minimum half of face area. Further, in the experiments, it was confirmed that the value of $P = 50$ guarantees the best results of recognition for test images having similar shadow coverage.

Experiments

Presented results come from two groups of experiments. In the first group of experiments the template base (training set) for FaReS contains 38 images from Subset 0 only (one image per subject). The testing set consists of all images from Subset 1 ÷ Subset 5, which gives 2414 test images, belonging to all 38 classes. The second group of experiments employs a database of FaReS, which includes 228 images from Subset 1 (6 images per class) and 2151 images from Subset 2 ÷ Subset 5 (from all 38 classes). All experiments were conducted on FaReS without extending its database with other images.

A model of the first group of experiments has the following form:

(9) $YaleB(38/1/2414)\{G\&Log/2DDCT: 192 \times 168 \rightarrow d \times d\}[F+MDC/L1/rank=1]$.

The second group of experiments has the following form:

(10) $YaleB(228/6/2151)\{G\&Log/2DDCT: 192 \times 168 \rightarrow d \times d\}[F+MDC/L1/rank=1]$.

The following denotations were used: $YaleB(38/1/2414)$ means a template set consisting of 38 images (one per class), while the test set contains 2414 images; $YaleB(228/6/2151)$ means 228 templates (6 per class), and 2151 test images; $G\&Log/2DDCT$ – means gamma correction/logarithm of template data and two-dimensional cosine transformation (i.e. transformations were taken only at the stage of template and test data projection); $192 \times 168 \rightarrow d \times d$ – denotes a size of original image and the dimensions of resulting matrices; $F+MDC/L1/rank=1$ – means features fusion procedure and minimum distance classifier with metric $L1$ based on first place holder ($rank=1$).

The graphical representation of the process dynamics and classification results in the framework of model (9) are shown in Fig. 5. Here, 135 out of 2414 images were classified incorrectly giving the total recognition rate equal to 94.4%. Solid lines show the borders between data sets.

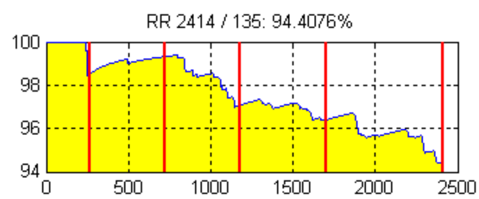

Fig.5. Classification accuracy for experiment described by (9)

Table 1 presents the precise classification results of test images within each data set for model (9). Initial component NK denotes the first element in the spectral matrix (starting from upper left corner of a matrix). For example for a parameter $NK=14$, we perform a zeroing of first 13 elements (corresponding to low frequency) of spectral matrix. Values of parameters d and NK were obtained during the modeling and testing of the classification algorithm.

Classification process and the result (dynamics) of test images in framework of model (10) are shown in Fig. 6. Table 2 below gives a closer look at the results of classification of images from Subset 2 ÷ Subset 5.

Table 1. The results of classification for model (9)

No. of templates – 38 (one per class)						
Subset No.	1	2	3	4	5	All
Parameter d	20	30	49	45	49	-
Parameter *NK*	1	1	1	1	14	-
No. of test images	263	456	455	526	714	**2414**
Misclassified images	4	1	30	27	73	**135**
Recognition rate	98.5	99.8	93.4	94.1	89.1	**94.4**

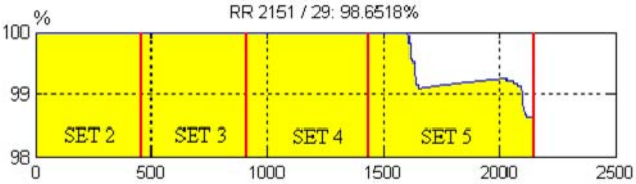

Fig.6. Classification accuracy for experiment described by (10)

Table 2. The results of classification for model (10)

No. of templates – 228 (6 per class)					
Subset No.	2	3	4	5	**All**
Parameter d	30	49	54	48	**-**
Parameter *NK*	1	1	1	14	**-**
No. of test images	456	455	526	714	**2151**
Misclassified images	0	0	0	29	**29**
Recognition rate	100	100	100	96.0	**98.7**

For templates database we used images from Subset 1, since classes no. 16-18 contain only 6 images each (and the rest of sets has 7 images per class). The first 6 images were used as templates for each class (228 total images). In each experiment only 29 images out of 2151 were not correctly classified. Hence, final result of classification reached 98.7%. Several misclassified images belonging to Subset 5 are presented in Fig. 7. The main reason of the misclassification is related to the low intensity and very high contrast (the values in the dark areas of images are close to zero, while the bright ones are close to the maximal value). It should be noticed that the mean intensities of presented images vary from 0 to 28.

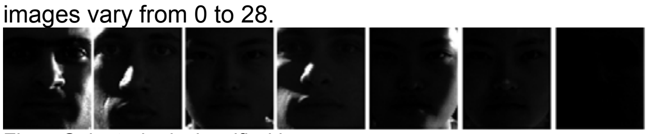
Fig.7. Selected misclassified images

Results shown in Fig. 6 and Table 2 correspond to the version of controlled classification of test images, and had no classification errors of FRR type (false rejection). In Fig. 8 we present an estimated boundaries for such errors. It can be seen that 74% of test images can be classified without errors (distance ≤0.94), while the distance higher than 0.94 results in errors of FRR type. This is a threshold value corresponding to the classification of the most difficult images from Subset 5.

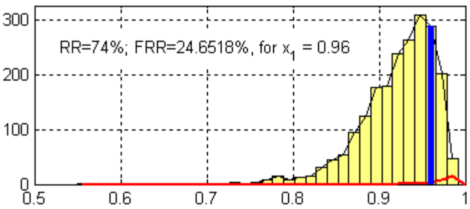

Fig.8. Estimation of FRR errors

Conclusions

It was shown that the 2DDCT method together with the brightness correction, fusion of features according to current mean value of brightness as well as removal of low frequency components of spectrum allow to achieve higher efficiency of recognition in case of facial portraits having illumination problems – flashes, shadows, and very low level of brightness.

The paper presents an exact model of conducted experiments, structure of corresponding FaReS, implementing its algorithm, and the results of tests executed on Yale B and Yale B+ databases. Obtained accuracy is better than ones presented in [1-10]. Moreover, the proposed algorithm is much more easy to describe and simple in implementation.

Wydanie publikacji zrealizowano przy udziale środków finansowych otrzymanych z budżetu Województwa Zachodniopomorskiego.

REFERENCES

[1] Chen W. et al., PCA and LDA in DCT domain, *Pattern Recognition Letters*, vol. 26 (2005), 2474–2482
[2] Chen W. et al., Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain,*IEEE Trans. Syst. Man Cybern. Part B*, vol. 36 (2006), No. 2, 458–466.
[3] Tan X., Triggs B., Preprocessing and Feature Sets for Robust Face Recognition, *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR'07 (2007), 1-8.
[4] Xiaohua Xie et. Al., Face Illumination Normalization on Large and Small Scale Features, *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR'08 Anchorage, AK (2008), 1 - 8.
[5] Abbas A. et. al., Illumination invariant face recognition in logarithm discrete cosine transform domain, IEEE *Internat. Conference of Image Processing*, ICIP'09, (2009) 4157-4160.
[6] Shao M., Wang Y., Joint Features for Face Recognition under Variable Illuminations, *Fifth International Conference on Image and Graphics*, ICIG'09 (2009), 922 – 927
[7] Liau HF., Isa D., New Illumination Compensation Method for Face Recognition, *International Journal of Computer and Network Security*, Vol. 2 (2010), No. 3, 308-321
[8] Hu H. et al., Lighting Aware Preprocessing for Face Recognition across Varying Illumination, *Lecture Notes in Computer Science 6312/ ECCV 2010*, Part II, 308–321
[9] Tripti G. et al., Comparative Analysis of various Illumination Normalization Techniques for Face Recognition, *International Journal of Computer Applications*, Vol. 28 (2011), No.9, 1-7
[10] Cao X. et al. Illumination invariant extraction for face recognition using neighboring wavelet coefficients, *Pattern Recognition* Vol. 45, (2012), 1299–1305
[11] Choi S. et al. Face recognition based on 2D images under illumination and pose variations, *Pattern Recognition Letters*, 32 (2011), 561–571
[12] The Extended Yale Face Database B http://vision.ucsd.edu/~leekc/ ExtYaleDatabase/ExtYaleB.html *(access 01.05.2012)*
[13] Kukharev G., Shchegoleva N., People recognition based on face images [in Russian]- SPB, CPBGETU (LETI), 2006, 176

***Authors**: Paweł Forczmański, PhD, West Pomeranian University of Technology, Szczecin, Faculty of Computer Science and Information Systems, ul. Żołnierska 52, 71-210 Szczecin, Poland, E-mail: pforczmanski@wi.zut.edu.pl; Prof. Georgy Kukharev, West Pomeranian University of Technology, Szczecin, Faculty of Computer Science and Information Systems, ul. Żołnierska 52, 71-210 Szczecin, Poland, E-mail: gkukharev@wi.zut.edu.pl; Nadezhda L'vovna Shchegoleva, PhD, Saint Petersburg State Electrotechnical University, ul. Prof. Popova 5, Saint Petersburg, 197376 Russia, E-mail: stil_hope@mail*