

Simple and Robust Facial Portraits Recognition under Variable Lighting Conditions Based on Two-Dimensional Orthogonal Transformations

Paweł Forczmański¹, Georgy Kukharev¹, and Nadezdha Shchegoleva²

¹ Faculty of Computer Science and Information Systems,
West Pomeranian University of Technology, Szczecin, Poland

{pforczmanski,gkukharev}@wi.zut.edu.pl

<http://pforczmanski.zut.edu.pl>

² Saint Petersburg State Electrotechnical University (LETI), Russia
stil_hope@mail.ru

Abstract. The paper addresses the problem of face recognition for images registered in variable lighting, which is common for real-world conditions. Presented algorithm is based on orthogonal transformation preceded by simple transformations comprising of equalization of brightness gradients, removal of spatial low frequency spectral components and fusion of spectral features depending on average pixels intensity. Two types of transformations: 2DDCT (two-dimensional Discrete Cosine Transform) and 2DKLT (two-dimensional Karhunen-Loeve Transform) were investigated in order to find the most optimal algorithm setup. The results of experiments conducted on Yale B and Yale B+ datasets show that a quite simple algorithm is capable of successful recognition without high computing power demand, as opposite to several more sophisticated methods presented recently.

Keywords: face recognition, illumination compensation, dimensionality reduction, DCT, PCA.

1 Introduction

The problem of facial portraits recognition has been attracting scientists interest for several recent years. Although it might has been solved in a satisfactory way, due to its extensive characteristic, it is still interesting. It is because most of the proposed solutions operate mainly in the very limited real-world conditions (i.e. controlled conditions of imaging, very restricted assumptions related to pose, orientation, expression etc.). The continuous research is also driven by practical implementations on low computing power devices like smartphones and tablets. Hence, there is still a need for robust solutions as simple as possible, yet without serious performance degradation.

One of the most vital issues in face recognition practice is a variation of light intensity in input images. In such case, we have to deal with two types of distortions perceived in the face area and its background. The first one is related

to the presence of local shadows, while the second one is linked to so called global shadows. Local shadows tend to change not only the form of elementary parts of face (eyes, nose and mouth) but also influence the boundaries of whole face area. On the other hand, global shadows significantly reduce the possibility to distinguish face areas from the background or even hide them completely.

Above problems highly degrade the accuracy of FaReS (Face Recognition System), thus this is one of the main reasons of constant interest of face recognition specialists [1–11]. Literature review shows, that the problem of variable lighting conditions is solved by one of just a few methods, namely 1. extending the FaReS database with a set of new images characterizing all distortions related to lighting problems, 2. transforming the images in order to find features that are independent of light characteristics, 3. reducing local intensity gradients or 4. transforming the images from spatial domain of into spectral one by means of models based on Eigenfaces, wavelets or cosine functions together with elimination of low-frequency components.

It should be emphasized, that in practice we have in our disposal only a limited training database, thus the implementation should employ FaReS base which is not extended with templates having all possible variants of shadows [4, 6, 9] as opposite to setups presented in [1–3, 5, 7, 8].

In order to successfully recognize images in difficult lighting conditions we often need to find their light-invariant representation. There are several approaches that combine different image characteristics, like Gabor features or Local Binary Patterns (LBP). Such an approach presented in [13] leads to the very complex multi-tier algorithm, which also requires a normalization of scales within different features.

One of the most popular approaches related to eigenbasis representation employs Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) [1–3], and Canonical Correlation Analysis (CCA) [6]. All of them often use Discrete Cosine Transform (DCT) as image preprocessing step.

There are also several approaches to the problem of illumination compensation involving wavelets, e.g. [10] or very complex models consisting of pose estimation and further illumination handling [11]. The latter requires a complex pose estimation (often performed manually), then performs edge filtering and PCA+LDA to estimate light direction. Finally, it compensates the illumination (providing it has properly recognized a direction) and extracts features using another LDA. Above approach is time-consuming and depends on many intermediate results, which in practice may be unpredictable.

The application of 2DDCT has been investigated and proved its high robustness at the stage of facial portraits recognition [1, 14]. In such methods, during feature extraction stage, the low frequency spectral components are often removed, as corresponding to shadows covering faces. The procedure in [2] involves the inverse transformation of images from spectral domain into spatial domain just before PCA step (which, according to authors' claims, increases the recognition accuracy). Similar approach [5], employs calculations related to PCA optimized by the Gramm matrix evaluation. Even though, presented results are

not better than average. The typical processing scheme employing DCT was also presented in [15]. The authors performed DCT on the whole image and kept first 49 coefficients later used in a standard PCA recognition scenario. Unfortunately, the authors fail to justify the selection of such number of spectral components. Moreover, the experiments were performed on a few small databases which do not give an objective knowledge about the method's robustness. Some local features derived from DCT were also employed in [16], however the performance of the method was evaluated on a very small number of images.

As it will be shown, above approaches do not give higher recognition accuracy in contrast to the method presented in our work. Although, the PCA used at the final stage of many of them is a bottleneck when it comes to fast database updates, it is often included in the processing pipeline. Hence, in this paper we focus on a comparison of two methods involving orthogonal transformations (PCA/KLT and DCT). It should be also noted, that most of approaches use simple 1D PCA, which requires eigen-decomposition of very large covariance matrices. In our approach we use 2D method [17] that employs smaller covariance matrices (which is many times faster and requires less memory storage).

This paper presents an algorithm aimed at lighting problems elimination in the aspect of facial portraits recognition, which is very simple in comparison to the earlier presented approaches, yet its accuracy is very similar. Its main features are: template database not containing images characterizing local and/or global shadows, adaptive preprocessing of original images aimed at reduction of intensity gradient (gamma correction and intensity logarithm), orthogonal transformation (2DDCT or 2DPCA/2DKLT) as the only instruments of transformation of original data into the low dimensional space, and finally, the adaptive combination of spectral features at the classification step.

2 Algorithm Details

2.1 Image Preprocessing

It should be noted that presented method is not fully invariant to rotations and shifts. However, there are several methods of face normalization that can be used. Hence, we assume that the input face is geometrically processed. On the other hand, it is evident that without brightness equalization the recognition of images taken under complex lighting conditions would be very difficult.

The methods of brightness enhancing can utilize gamma correction or logarithmic correction which compensate light intensity for properties of human vision, in order to maximize the use of bandwidth relative to how humans perceive light [1–9]. However, it may influence such parameters of input image as mean value, contrast, local brightness and boundaries of shadows. Moreover, it may partially change the form and characteristics of image, which can have a negative influence on the recognition accuracy. Figure 1 presents images taken under variable lighting conditions and the results of applying one of two enhancing procedures: gamma correction (G) and brightness logarithm (Log). After such operations we

can still observe some distortions (like shadows, bright spots, low contrast and noise). Despite this fact, the boundaries of all face parts can be easily detected, and its anthropometric parameters may be successfully explored. It can be also seen that above operations unveil different parts of face, which are originally hidden in the shadows, hence it may increase the FaReS accuracy. In contrary to the works presented earlier, in our approach we propose a fusion of features obtained after gamma correction and logarithmic transform for images with extreme lighting conditions. According to our observations, the combination of these two types of preprocessing stages outperforms single one especially when an image is very dark. Hence, the further increase in the recognition accuracy is caused by the adaptation to the image brightness.



Fig. 1. Gamma correction (G) and intensity logarithm (Log) applied to original images

In many previous works additional stages of processing were based on PCA/KLT approach. Our experiments described recently [18] and in the further part of this paper proved, that using much simpler approach, namely DCT, makes it possible to obtain good recognition results. The basis functions of cosine transform are optimal (in terms 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 decided to 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 [14]. However, for the comparison purposes, we show the results of experiments involving KLT.

2.2 Recognition Stage

Developed algorithm for image of size $M \times N$ consists of five following stages, and is common for both 2DDCT and 2DPCA/2DKLT methods (see Fig. 2):

1. Two projection matrices are build: F_1 and F_2 of sizes $d_1 \times M$ and $N \times d_2$, respectively ($d_1 \ll M$ and $d_2 \ll N$; $d_1 \neq d_2$ in general case). In case of 2DDCT this stage is equal to selecting proper functions from the cosine basis, while in case of 2DPCA/2DKLT, such functions are calculated by means of 2DPCA/2DKLT applied for training images only [17]. The step related to 2DPCA/2DKLT requires extra computations, and from the practical point of view, increases the computational complexity of the whole approach.

2. Transformation (2DDCT or 2DPCA/2DKLT, respectively) for templates I with preliminary processing, described above, resulting a database of processed templates \tilde{I} . We define both two-dimensional linear transformations in a matrix form as the following multiplication: $Y = F_1 \tilde{I} F_2$, where Y - transformation result; F_1 and F_2 - projection matrices of size $d_1 \times M$ and $N \times d_2$, Note, that the projection of original images of size $M \times N$ into new feature space, results in a new representation with spectral matrices Y of size $d_1 \times d_2 \ll MN$.
3. Gamma correction (G) or gamma correction followed by logarithmic correction (G&Log) performed on input image. The preprocessing type depends on average image brightness J . If values of J is lower than a certain threshold P , then both optional steps are executed, otherwise only one step i.e. gamma correction or logarithmic correction is employed (the performance of both variants is similar).
4. Transformation of preprocessing results into a new feature space based on 2DDCT/2DKLT. Projection result consists of spectral matrices (precisely, one or two matrices) of size $d_1 \times d_2 \ll MN$. 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 $d = \min(d_1, d_2)$.
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 L_1 . Test image is assigned to a class (one of known classes) having a minimum distance.

The problem of recognizing extremely dark images is solved through the following (additional) processing. Before the classification stage, mean brightness J of input image is calculated and if it is lower than a predefined threshold $J < P$, then the projection results (step 4 of the algorithm) are fused in a one common vector, which is compared to respective collection of features stored in the database. In presented experiments, a value of P was set to 110, as corresponding to shadows covering minimum half of face area.

3 Experiments

3.1 Characteristics of Experimental Data

Most of the reported experiments in the area of facial portrait recognition in the presence of variable lighting are conducted on Yale database (including original Yale B images and its extension called Yale B+) [12]. It features a quite wide diversity of illumination variations and the proper registration of images, namely they all are cropped and normalized in such way, that the eyes positions are constant in all images across all subjects. It allows to investigate the illumination influence without the overhead of geometrical transformation.

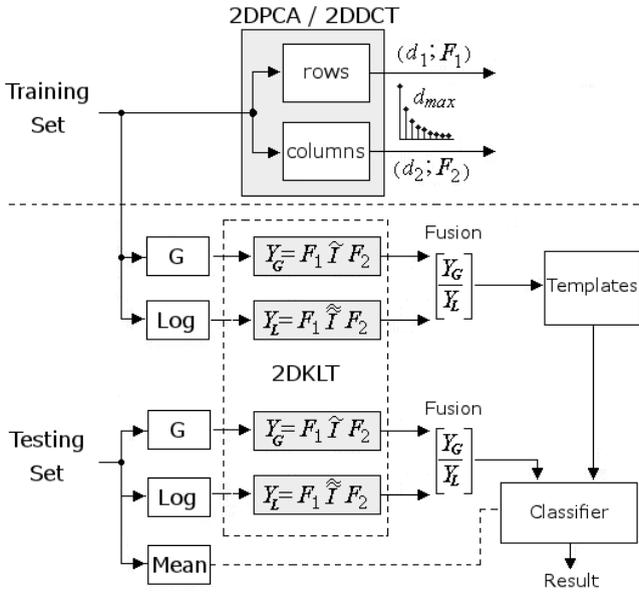


Fig. 2. Structure of a recognition system based on 2DPCA/2DDCT

In our 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 eliminated since they cannot be read from files published on web site [12]). The grayscale images are stored in matrices of 192×168 pixels. They are divided into 6 sets, labeled Subset 0 – Subset 5, respectively. In Fig. 3 one can see exemplary images coming from this dataset. Images in Subset 0 have no blinks nor shadows and are evenly illuminated. Images in Subset 1 – Subset 5 were obtained by modeling the spatial movement of a light source, thus they represent 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 challenging are images from Subset 3 – Subset 5.



Fig. 3. Selected images from Yale set: Subset 0 (left) and Subsets 1-5 (right)

3.2 Results

The experiments employed two scenarios of dividing the database: the first one employed a training set consisting of 38 images (one image per each class) and 2414 testing images, while the second one consisted of 228 training images (6 images per class) from Subset 1 and 2151 testing images from Subset 2 – Subset 5 (from all 38 classes). All experiments were conducted without extending the database with any other preprocessed images.

We selected G&Log approach as a compromise between complexity/accuracy and investigated the influence of training set selection strategy and transformation type on the recognition accuracy. We investigated the influence of the number of spectral components on the recognition accuracy (results in Tab. 1). In most cases we used 20–40 first spectral components, thus they are strongly related to illumination characteristics of the input image. For images taken from Subset 5, we selected components that are further from the origin, which makes it possible to remove any artifacts related to directional illumination.

Table 1. The results of classification involving 2DDCT and 2DPCA/2DKLT

Number of templates	38 (one image per class)						228 (6 images per class)				
Subset No.	1	2	3	4	5	All	2	3	4	5	All
Number of test images	263	456	455	526	714	2414	456	455	526	714	2151
Transformation	2DDCT										
Misclassified images	3	1	30	27	69	130	0	0	1	16	17
Recognition rate [%]	98.9	99.8	93.4	94.9	90.3	94.6	100	100	99.8	97.8	99.2
Transformation	2DPCA/2DKLT										
Misclassified images	4	1	31	43	91	170	0	0	4	22	26
Recognition rate [%]	98.5	99.8	93.2	91.8	87.2	93.0	100	100	99.2	96.9	98.8

The experiments exposed the problem of recognizing images taken at a very low level of illumination. Several misclassified images are presented in Fig. 4 (most of them belong to Subset 5). As it can be seen, these images have a very low intensity level and a very high contrast. The proposed preprocessing scheme does not increase the visual quality of such images since they contain hardly no information. Mean intensities of presented images vary from 0 to 28. This observation may be a valuable hint for developing a practical implementation in which images with such low brightness level should be rejected as unusable.

3.3 Comparison with State-of-the-Art

It should be noted, that there are many recent methods that use DCT and characterize good recognition accuracy. One of them [1] employs DCT+PCA. The reported accuracy is 94%, which is lower than ours. The most similar approach [2] includes logarithm operator, DCT and low-frequency components extraction. The reported accuracy (98.3%) is a little bit lower than ours. The perfect recognition (100%) is reported in [24]. However, it works on extracted facial parts and

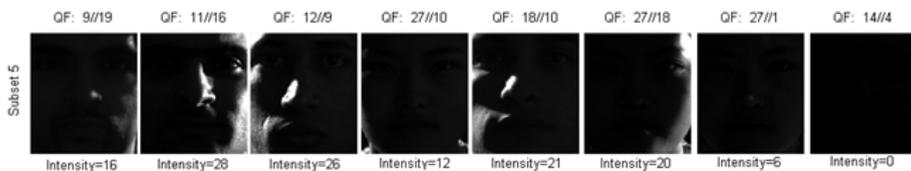


Fig. 4. Selected misclassified images

the increase in the recognition accuracy was achieved through the independent recognition of all facial parts and a highly-complex processing (reduction of DCT components, cross-correlation, PCA, k-NN etc.).

The face recognition with illumination suppression involving other popular techniques is provided below. The experiments were performed on Yale B and Yale B+ datasets. We focus on recognition rates for the following methods: Single Scale Retinex algorithm (SSR) [19], Logarithmic Total Variance technique (LTV) [20], Wavelet-Denoising-based method (WD) [13] and Nuisance Attribute Projection (NAP) [21]. All above techniques were implemented without any pre- or post-processing (e.g., histogram equalization) [21]. The computational expenses comparison of those methods shows that the proposed method is less complex. NAP [21] requires a projection matrix for all the images in the database. When the database changes, the projection should be re-build. Moreover, it requires an eigenvectors decomposition which is itself very time-consuming task. The LTV [20] also is not a trivial one. It features an edge detection stage (convolution filtering) at different scales and again, PCA, at the stage of features reduction. The SR [19] is a global approach and the results of directional lighting compensation are much worse than any other propositions. The last of the methods addressed in the paper [13] requires fast wavelet transform but again, PCA at the final stage slows the whole method down.

The results of the accuracy comparison (see Tab. 2) suggest that our technique gives competitive results when compared to other state-of-the-art methods.

Table 2. Comparison of the performance [%] of other state-of-the-art methods

Subset No.	Our Method	NAP	SSR	LTV	WD
2	100	100	100	100	100
3	100	100	99.2	100	100
4	99.8	99.3	82.9	99.3	98.6
5	97.8	96.8	81.1	99.5	99.5

4 Summary

The article presents a simple, yet effective method of recognizing facial portraits taken under difficult lighting conditions. There is still a gap between simple methods (fast and easy to implement) and sophisticated ones (that require complex

computations and tuning). Our approach involves simple preprocessing stages (gamma correction and logarithmic correction) and feature extraction by means of two-dimensional orthogonal transformations (2DDCT and 2DPCA/2DKLT). They are handled by many hardware systems that work on fixed-point arithmetic. Here, we propose an adaptive fusion of features obtained after gamma correction and logarithmic transform for images with extreme lighting conditions which increases overall recognition accuracy. During experiments performed on Yale datasets 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 with illumination problems. The accuracy of the developed algorithm is better than ones presented in [1–10, 13, 19–21]. Moreover, it is easier to describe and simple in implementation and does not require any other pre-processing (like downsampling [4]).

Another interesting observation is related to the differences between 2DDCT and 2DPCA/2DKLT as the main transformation stage. Despite theoretically better adaptation of basis functions (2DPCA), in our experiments, 2DDCT components perform with slightly better efficiency. Its higher efficiency is caused by the fact, that 2DPCA/2DKLT requires much wider representation of training data to form an optimal eigenbasis. In the performed experiments the basis was calculated using quite low number of images which resulted in worse adaptation to the testing set. In contrary, 2DDCT components in such case give a much better basis. This is especially visible when the basis is calculated for images without directional illumination, strong shadows or low brightness. We showed, that Eigenfaces method is not always the best choice when it comes to appearance-based recognition.

References

1. Chen, W., Er, M.J., Wu, S.: PCA and LDA in DCT domain. *Pattern Recognition Letters* 26, 2474–2482 (2005)
2. Chen, W., Er, M.J., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans. Syst. Man Cybern. Part B* 36(2), 458–466 (2006)
3. Tan, X., Triggs, B.: Preprocessing and Feature Sets for Robust Face Recognition. In: *IEEE Conf. on Computer Vision and Pattern Recognition, CVPR 2007*, pp. 1–8 (2007)
4. Xie, X., Zheng, W.-S., Lai, J., Yuen, P.C.: Face Illumination Normalization on Large and Small Scale Features. In: *IEEE Conf. on Computer Vision and Pattern Recognition, CVPR 2008 Anchorage*, pp. 1–8 (2008)
5. Abbas, A., Khalil, M.I., AbdelHay, S., Fahmy, H.M.A.: Illumination invariant face recognition in logarithm discrete cosine transform domain. In: *IEEE Inter. Conf. of Image Processing, ICIP 2009*, pp. 4157–4160 (2009)
6. Shao, M., Wang, Y.: Joint Features for Face Recognition under Variable Illuminations. In: *Fifth Inter. Conf. on Image and Graphics, ICIG 2009*, pp. 922–927 (2009)

7. Liao, H.F., Isa, D.: New Illumination Compensation Method for Face Recognition. *Inter. Journal of Computer and Network Security* 2(3), 308–321 (2010)
8. Han, H., Shan, S., Qing, L., Chen, X., Gao, W.: Lighting Aware Preprocessing for Face Recognition across Varying Illumination. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) *ECCV 2010, Part II. LNCS*, vol. 6312, pp. 308–321. Springer, Heidelberg (2010)
9. Goel, T., Nehra, V., Vishwakarma, V.P.: Comparative Analysis of various Illumination Normalization Techniques for Face Recognition. *Inter. Journal of Computer Applications* 28(9), 1–7 (2011)
10. Cao, X., Shen, W., Yu, L.G., Wang, W.L., Yang, J.Y., Zhang, Z.W.: Illumination invariant extraction for face recognition using neighboring wavelet coefficients. *Pattern Recognition* 45, 1299–1305 (2012)
11. Choi, S., Choi, C.-H., Kwak, N.: Face recognition based on 2D images under illumination and pose variations. *Pattern Recognition Letters* 32, 561–571 (2011)
12. The Extended Yale Face Database B, <http://vision.ucsd.edu/~lseekc/ExtYaleDatabase/ExtYaleB.html> (accessed May 01, 2012)
13. Zhang, T., Fang, B., Tang, Y.Y., Shang, Z., Li, D., Lang, F.: Multiscale facial structure representation for face recognition under varying illumination. *Pattern Recognition* 42(2), 251–258 (2009)
14. Forczmański, P., Kukharev, G.: Comparative analysis of simple facial features extractors. *Journal of Real Time Image Processing* 1(4), 239–255 (2007)
15. Hafed, Z.M., Levine, M.D.: Face Recognition Using the Discrete Cosine Transform. *International Journal of Computer Vision* 43(2), 167–188 (2001)
16. Schwerin, B., Paliwal, K.: Local-DCT features for facial recognition. In: 2nd Inter. Conf. on Signal Processing and Communication Systems, ICSPCS 2008, pp. 1–6 (2008)
17. Kukharev, G., Forczmański, P.: Facial images dimensionality reduction and recognition by means of 2DKLT. *Machine Graphics & Vision* 16(3/4), 401–425 (2007)
18. Forczmański, P., Kukharev, G., Shchegoleva, N.: An algorithm of face recognition under difficult lighting conditions. *Przegląd Elektrotechniczny (Electrical Review)* 10b, 201–205 (2012)
19. Jobson, J., Rahman, Z., Woodell, G.A.: Properties and performance of a Center/Surround Retinex. *Trans. on Image Processing* 6(3), 451–462 (1997)
20. Chen, T., Yin, W., Zhou, X.S., Comaniciu, D., Huang, T.S.: Total variation models for variable lighting face recognition. *TPAMI* 28(9), 1519–1524 (2006)
21. Struc, V., Vesnicer, B., Mihelic, F., Pavesic, N.: Removing illumination artifacts from face images using the nuisance attribute projection. In: *ICASSP 2010*, pp. 846–849 (2010)
22. Zhiming, L., Chengjun, L.: Fusion of color, local spatial and global frequency information for face recognition. *Pattern Recognition* 43(8), 2882–2890 (2010)
23. Akrouf, S., Sehili, M.A., Chakhchoukh, A., Mostefai, M., Youssef, C.: Face Recognition Using: PCA and DCT. In: *Proceedings of the 2009 Fifth International Conference on MEMS NANO, Smart Systems, ICMENS 2009*, pp. 15–19 (2009)
24. Vishwakarma, V.P., Pandey, S., Gupta, M.N.: An Illumination Invariant Accurate Face Recognition with Down Scaling of DCT Coefficients. *Journal of Computing and Information Technology - CIT* 18(1), 53–67 (2010)