Recognition of Occluded Faces Based on Multi-subspace Classification

Pawel Forczmański, Piotr Łabędź

Recognition of Occluded Faces Based on Multi-subspace Classification

Paweł Forczmański and Piotr Łabędź

West Pomeranian University of Technology, Szczecin,
Faculty of Computer Science and Information Systems,
ul. Zolnierska 52, 71-210 Szczecin, Poland,
{pforczmanski, plabedz}@wi.zut.edu.pl,
WWW home page: http://pforczmanski.zut.edu.pl

Abstract. In the paper we investigate a problem of face recognition in uncontrolled environment – distorted by occlusion, shadows and other local modifications. Such problems are very common for real-world conditions, thus the presented algorithm allows to eliminate them. It is based on dimensionality reduction approach (two-dimensional Karhunen-Loéve Transform) and distance-based classification. We use simple transformations involving face normalization and individual facial regions extraction as a pre-processing. Then, we perform independent recognition of extracted facial regions and combine the results in order to make a final classification. The results of experiments conducted on images taken from 9 publicly available datasets show that a quite simple algorithm is capable of successful recognition without high computing power requirements, as opposite to more sophisticated methods presented in the literature. As it was proved, the presented approach gives significantly better efficiency than a whole image-based recognition.

Keywords: face recognition, subspace method, dimensionality reduction, PCA

1 Introduction

The problem of automatic face recognition continuously attracts scientists interest. The review of literature and closer look at practical implementations show that although it might has been successfully solved, due to its broad characteristic, it is still interesting. This constant interest is motivated by the observation that most of the proposed solutions operate mainly in the very narrow real-world conditions (i.e. in the controlled conditions of imaging, with strictly defined illumination, orientation, pose, expression etc.).

The continuous research in this area is influenced also by the fact that facial recognition algorithms are to be implemented in the devices having low processing power and very limited resources like smartphones and tablets. Hence, robust solutions as simple as possible, yet without suffering from significant performance degradation are still needed.

One of the most important problems in face recognition is a low quality of input data processed by a typical Face Recognition System (FaReS). It should be noted, that the quality is often degraded by two main types of distortions observed in the face area: variable lighting and occlusions/incompleteness of face portrait.

Several examples of face occlusions or similar local distortions that can be found in practice are presented in Fig.1. As it can be seen, they include full occlusion of different face areas, partial occlusion by an independent object, local intensity changes or total information loss in a limited region of face. Those problems highly degrade the performance of FaReS, thus this is the main reason of lasting interest of researchers [1].

Fig. 1. Real distortion examples in the face area

It should be also emphasized, that in practice we have an access to a limited training database, thus we have to solve the facial portrait recognition without extending FaReS database with images having all possible variants of global and/or local distortions (in contrast to setups defined in [1]). The literature review shows also, that recognizing images under occlusion requires a careful selection of invariant features. Apart from many complex geometry-based approaches, there are several simple appearance-based methods that combine a few elementary features to describe an image, like Gabor Features[2] or LBP Features [3]. However they lead to the very complex multi-tier algorithms, which is unusable in low-performance devices.

One of the most popular approaches to face recognition employs Principal Component Analysis (PCA) also known as Karhunen-Loeve Transform (KLT) [1, 4, 6–9]. There has been many scientific works published in the recent years related to the data dimensionality reduction by means of PCA/KLT, as well as their application to the face recognition problem taking into account the two-dimensional characteristics of images. Such algorithms have got different names i.e. 2DPCA [10, 11] or MatPCA [12]. However, most of them are associated with the processing of two-dimensional data by means of data reduction algorithms performed along one dimension only or by dividing an image into smaller parts

and reducing them with help of classical PCA. Moreover, those methods operate on fully visible facial portraits not considering a problem of occlusion and local distortions.

The literature review shows that there is a few attempts to deal with above problems, however they are still not fully solved. An interesting approach was presented by the authors of [13]. They address a problem of reconstructing images using so called Fast-Robust PCA. The experimental part proves that it is possible to rebuild missing pixels in images of known class, however it is computationally expensive. The improved method presented in [14] proved that it is possible to reconstruct distorted pixels in facial portraits using a modified (faster) PCA approach. That method makes it possible to reconstruct occluded regions of the face using Fast Weighted Principal Component Analysis (FW-PCA). The occluded regions are detected and recursively updated. Unfortunately, the authors do not address the problem of recognition accuracy. The extension of classical PCA, called Lophoscopic PCA, aimed at recognition of images under occlusion was presented in [15]. The authors covered only small parts of faces (one part a time) which do not represent real-life conditions. Moreover, the main drawback of above method is a complex and time-consuming calculations (six times slower than a classical PCA).

Another interesting algorithm was presented in [16]. The authors propose an approach which consists of first detecting the presence of scarf/sunglasses and then processing the non-occluded facial regions only. The occlusion detection problem is solved using Gabor wavelets, PCA and Support Vector Machines (SVM), while the recognition of the non-occluded facial part is performed using block-based Local Binary Patterns (LBP). The main drawback in this case is the assumption about the presence of sunglasses or the scarf, not taking into account other distortions. Moreover, LBP are not the most optimal face recognition approach. On more interesting approach was presented in [17], however that solution is devoted to a removal of one specific object, namely glasses.

The general disadvantage of above methods is the computational complexity coming from the iterative nature of calculating a reconstructed image or the need of detecting occluded regions. Thus, it is a major drawback when it comes to robust and fast face recognition.

In this paper we focus on a method involving dimensionality reduction approach operating in two-dimensional domain of KLT, as it proved to be very efficient, yet not very computationally expensive, especially in case of currently available computing power. We present an algorithm of solving occlusion problem in the aspect of facial portraits recognition, which is much simpler in comparison to the above presented approaches, yet its efficiency is very similar. It employs a template database that does not include images with local distortions. At the stage of preprocessing it uses face normalization and facial regions extraction. At the stage of feature extraction it uses 2DPCA/2DKLT as the only instruments of transformation of original data into the low dimensional space [18]. Finally, obtained feature vectors are used at classification step (using simple distance metrics).

2   Processing Algorithm

The main idea behind the recognition algorithm is the observation, that in most cases faces are occluded only to some extent. If a human being is able to recognize such an occluded portrait, a computer also can do it. So we assume the following. If we can successfully recognize *most* of facial parts, then we can decide about the recognition of the whole face.

Facial portraits have different geometrical characteristics, hence it is important to normalize them. Here we assume to have *en face* portraits. The pre-processing of such images includes orientation normalization (based on eyes positions), face cropping and scale standardization. For the purpose of experiments presented in this paper, the eyes positions were marked manually, hoverer it seems that using some sort of automatic approach (i.e. Viola-Jones detector) may give similar results. The intermediate results of pre-processing are presented in Fig.2.

Then, the distances between salient points are used to crop individual face parts: forehead, eyes, nose and mouth/chin. Such distances used at this stage are derived from the distance between eyes (a in Fig.2): $b = 1.5 * a$, $c = 1.8 * a$, and $d = 2 * a$.
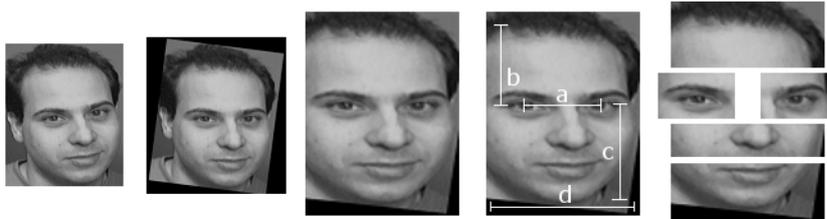
Fig. 2. Processing stages: original image, orientation normalization, scale standardization, salient point detection, facial areas extraction

The recognition algorithm consists of several stages: image pre-processing, facial regions extraction, dimensionality reduction, and recognition. The last stage includes parallel classification in multiple sub-spaces and final voting. In our proposition we divide a face into five, non-overlapping parts and perform an independent recognition on these parts. Such decomposition is motivated by a natural structure of face with five clearly distinguishable areas. As a result, we get several recognition results related to forehead, eyes, nose and mouth/chin areas. Each part belongs to a certain individual enrolled in the database. When we perform a voting on those results we get the final recognition result based on majority of votes.

The recognition scheme is presented in Fig.3. The *Facial Region Extraction* is a block responsible for face region extraction, *2DKLT* is a dimensionality reduction stage, C is a comparator (distance metrics) and *Result* blocks represent individual recognition results.
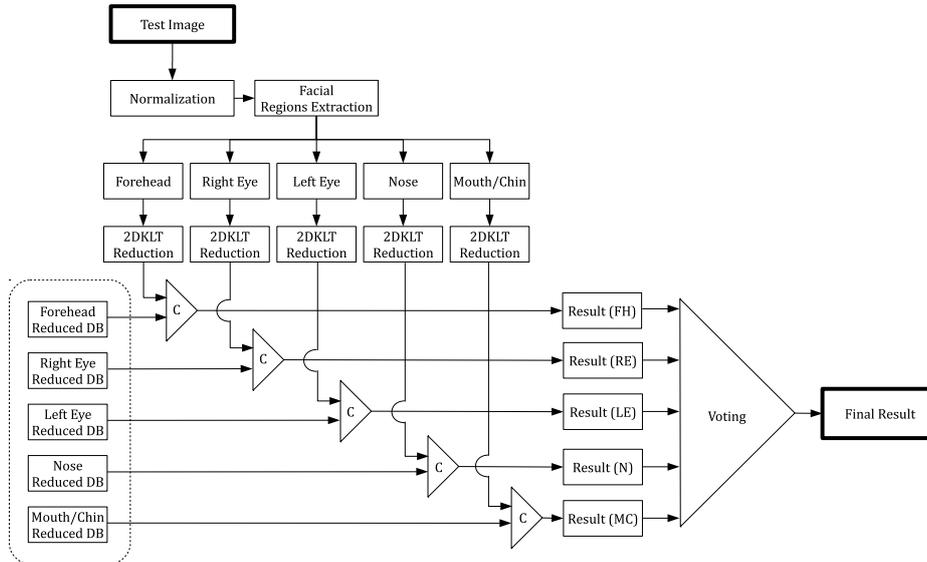
Fig. 3. Recognition scheme

Recognition of images is performed in subspaces related to each individual facial region. While the database images were registered in a precise manner (eyes positions were marked manually), test images are cropped in an approximate manner (since there is no guarantee, the eyes are visible). If the eyes are covered or impossible to localize, some different salient point detector may be used in order to calculate appropriate proportions (i.e. [2]).

Each region forms a database with reduced features and an information about the individual (*owner* of this particular part). We employ two-dimensional variant of well-known PCA/KLT approach as it makes it possible to obtain a compact representation of features [18]. As a distance metrics D_k for individual regions ($k = 1, 2, \ldots 5$) we use the Chebyshev distance defined on a vector space where the distance between two feature vectors P and Q is the greatest of their differences along any (i-th) coordinate dimension:

$$D_k(P, Q) = \max_i(|p_i - q_i|). \tag{1}$$

Then we use a 1-Nearest Neighbor approach to find a class center (average vector for all elements in each class), which is the closest one. Finally we perform a voting on the individual results $D_1, D_2, \ldots D_5$. We use simply the mode average of all votes, hence the majority indicates the class of the input image. Such an approach results in a fast and reliable recognition without high computational overhead.

Another interesting and practically important feature of presented approach is the possibility to compute a level of confidence. It is evaluated as a number of unanimous votes to the total number of votes. For example, if the voting results

in 3 votes indicating the same class (the rest 2 votes indicate different classes), then the level of confidence is equal to 3/5. The minimal acceptable confidence level is equal to 2/5 (40%). Hence we can evaluate the quality of recognition and give a hint to the human operator about the system's performance.

In the voting scheme, there can occur a special case, when there is an equality situation among the votes, i.e. two of the regions say one identity, and two other regions say another identity. It should be noted, that we do not weight the individual results (we do not consider one part of a face be more important than another). Hence, the system results in a identity that is associated with face parts placed in the upper part of the head. In the future, some more sophisticated solutions could be employed, such as presented in [19].

3 Experiments

First, it should be noted, that there is no publicly available large database aimed at testing face recognition under occlusion. One possibility is to use AR database [20] which features frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). However it does not incorporate images with occlusions of different type, i.e. by independent objects, strong local shadows etc. Thus, the experiments were performed on a self-created database containing images taken from 9 officially available sets: Olivetti Research Lab (AT&T) [21]: 400 images (40 individuals, 10 image per class); Bio ID [22]: 1520 images (23 individuals); Nottingham DB [23]: 465 images (71 individuals); JAFFE DB [24]: 213 images (10 individuals); Georgia Tech Face Database [25]: 750 images (50 individuals); Aberdeen DB [26]: 690 images (90 individuals); Caltech Faces 1999 [27]: 450 images (27 individuals); Pain Expression [28]: 599 images (23 individuals); FEI Face Database [29]: 2800 images (200 individuals);

The final database consist of *en face* portraits only, thus many original images were excluded. All images were converted to grayscale and re-scaled to 160×120 pixels. The dimensions of each facial region extracted from facial portraits are presented in Tab.1. The total number of images is equal to 3445 divided into 365 classes. Each class contains $4 - 24$ images (depending on the source database). The database was split into learning and testing datasest in a random way, so the learning set contains 3106 and the testing set 1695 images (5 distortion variants for each of 339 base images, respectively).

Table 1. Dimensions of images representing face regions and mean images

Face part	width	height	% of area
Forehead	120	53	33
Left eye	60	36	11
Right eye	60	36	11
Nose	120	26	17
Mouth/Chin	120	45	28

   The testing images were altered (using simulated occlusion) and divided into
5 categories, related to the distortion type. Each category is characterized by
the area of distortion, the form of distortion and the mean information loss (in
terms of the number of pixels changed in comparison to the original image). The
information loss/distortion of each part of a face is presented in Tab. 2. The
following categories of distortions are included in the database (the examples of
simulated images representing above distortions were presented in Fig. 4):

(A) Reference images (with no distortions);
(B) Eyes covered; mean information loss: 16%;
(C) Face area covered with randomly placed object; mean information loss: 13%;
(D) Face area with lower intensity; mean information loss: 28 %;
(E) Randomly covered face area; mean information loss: 36 %.

Fig. 4. Distortion examples used in the experiments (B, C, D, and E, respectively)

Table 2. Mean information loss in terms of percentage of changed (lost) pixels in each
part in comparison to the each part area and the whole image, respectively

Face part	Distortion category			
	B	C	D	E
Forehead	0 (0)	1.7 (0.6)	9.2 (3.1)	39 (12.9)
Left eye	72.3 (8.1)	4.1 (0.5)	13.9 (1.6)	22 (2.5)
Right eye	71 (8)	22.7 (2.6)	48.7 (5.5)	20 (2.3)
Nose	0 (0)	22.6 (3.7)	44.3 (7.2)	33.8 (5.5)
Mouth/Chin	0 (0)	20.5 (5.8)	40.2 (11.3)	44.6 (12.5)

We performed recognition employing 16, 100 and 400 principal components per
image region (80, 500 and 2000 PCs per face image, respectively). In order to
evaluate the performance of our method, we additionally performed a recogni-
tion of whole images employing 100 and 400 PCs. The results of recognition of
sample images are presented in Fig.5, where *query* means testing image, *recon-
struction* represents image composed of individually recognized regions, *result(1)*
represents our algorithms while *result(2)* shows recognition by whole image ap-
proach. The group of pictures on the left hand side show the results where the
accuracies of region-based and whole image-based recognition are equal. On the

right hand side of the figure, the results of region-based approach are superior to whole image-based approach.

Fig. 5. Sample recognition results

The comparative results are presented in Fig.6. As it can be seen, the recognition of undistorted faces involving whole images - case (A) - gives the highest accuracy (97.6% vs. 95%). In case of portraits with eyes covered with black rectangle - case (B) - presented method of recognition provides lower rate (83.8% vs. 88.5%). Images covered with independent object - case (C) - give the highest recognition rate for whole-image approach (93.8% vs. 90.9%). The rest two cases (D) and (E), where the information loss is the highest show the true potential of presented method. The recognition rates for these distortion types and region-based processing are equal to 63.1% vs. 56.9% and 44.2%, respectively.

4   Summary

We presented a novel algorithm of recognizing facial portraits under local distortions (occlusion, shadows, intensity change). It incorporates face normalization, extraction of individual face regions, dimensionality reduction and classification
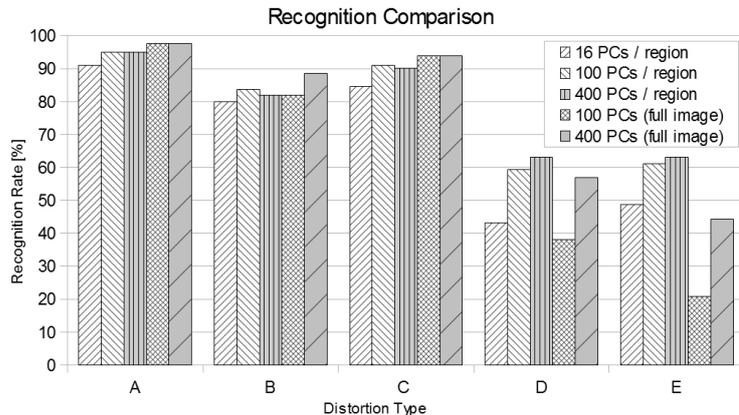
Fig. 6. Comparison of recognition performance for different variants of methods

in multiple sub-spaces. The final result is obtained by a voting with a majority principle. The results of experiments performed on large database show the superiority of developed method over traditional approach, mainly in case of large information loss. In case of images not influenced by above distortions, the recognition rate is similar. Although, we did not investigated the influence of low light intensity or degraded image quality (in case of low resolution and low sensitivity cameras), we think the presented method based on PCA is able to cope with such problems (as it has been proved in the literature many times before). In future, another extension of 2DPCA may be employed, namely the processing of colour images [30], in order to make the whole approach more robust and capable of distinguishing between facial and non-facial features.

References

1. Tan, X., Triggs, B.: Preprocessing and Feature Sets for Robust Face Recognition. IEEE Conf. on Computer Vision and Pattern Recognition - CVPR'07. (2007) 1–8
2. Choras, M., Andrysiak, T.: Symmetry-based Salient Points Detection in Face Images. in L. Rutkowski et al. (Eds): Artificial Intelligence and Soft Computing - ICAISC 2006. LNCS 4029. Springer-Verlag (2006) 758–767
3. 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. Vol. 42. Issue 2 (2009) 251–258
4. Chen, W., Er, M.J., Wu, S.: PCA and LDA in DCT domain. Pattern Recognition Letters 26 (2005) 2474–2482
5. Choi, S., Choi C.-H., Kwak N.: Face recognition based on 2D images under illumination and pose variations. Pattern Recognition Letters 32 (2011) 561–571
6. Sirovich, I., Kirby, M.: Low-dimensional procedure for the characterization of human faces, Journal of Opt. Soc. Am. 4 (1987) 519–524
7. Turk, M., Pentland, A.: Eigenfaces for Recognition. Journal of Cognitive Neurosicence 3(1) (1991) 71–86

8. Swets, D.L., Weng, J.: Using Discriminant Eigenfeatures for Image Retrieval. IEEE Trans. Pattern Analysis and Machine Intelligence 18 (1996) 831–836
9. Tsapatsoulis, N., Alexopoulos, V., Kollias, S.: A Vector Based Approximation of KLT and Its Application to Face Recognition. Proceedings of The IX European Signal Processing Conference EUSIPCO-98. Island of Rhodes, Greece (1998) 1581–1584
10. Yang, J., Zhang, D., Frangi, A.F., Yang J.-Y.: Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 26(1) (2004) 131–137
11. Nagabhushan, P., Guru, D. S., Shekar, B. H.: Visual learning and recognition of 3D objects using two-dimensional principal component analysis: A robust and an efficient approach. Pattern Recognition 39(4) 721–725
12. Chen, S., Zhu, Y., Zhang, D., Yang, J.-Y.: Feature extraction approaches based on matrix pattern: MatPCA and MatFLDA. PRL 26 (2005) 1157–1167
13. Storer, M., Roth, P.M., Urschler, M., Bischof, H.: Fast-Robust PCA. Proc. 16th Scandinavian Conference on Image Analysis (2009) 430–439
14. Hosoi, T., Nagashima, S., Ito, K., Aoki, T.: Reconstructing occluded regions using fast weighted PCA.Image Processing (ICIP). 19th IEEE International Conference on. (2012) 1729-1732
15. Tarrés, F., Rama, A., Torres, L.: A Novel Method for Face Recognition under Partial Occlusion or Facial Expression Variations. 47th International Symposium ELMAR-2005. Multimedia Systems and Applications. Zadar (2005) 163–166
16. Min, R., Hadid, A., Dugelay, J.: Improving the recognition of faces occluded by facial accessories, Automatic Face Gesture Recognition and Workshops (FG 2011), IEEE International Conference on (2011) 442–447
17. Park, J.-S., Oh, Y.H., Ahn, S.C., Lee S.-W.: Glasses removal from facial image using recursive error compensation, Pattern Analysis and Machine Intelligence, IEEE Transactions on 27(5) (2005) 805–811
18. Kukharev, G., Forczmański, P.: Data Dimensionality Reduction for Face Recognition. Machine Graphics and Vision, 13(1/2) (2004) 99–122
19. Gokberk B., Dutagaci H., Ulas A., Akarun L., Sankur B.: Representation plurality and fusion for 3-D face recognition. Trans. on Systems Man and Cybernetics, Part B. 38(1) (2008) 155–173
20. Martinez, A.M., Benavente R.: The AR Face Database. CVC Technical Report #24, June (1998)
21. AT&T Laboratories Cambridge. Database of Faces. http://www.cl.cam.ac.uk/ research/ dtg/attarchive/facedatabase.html
22. BioID-Technology Research. The BioID Face Database. http://www.bioid.com/downloads/facedb/index.php
23. Nottingham DB. http://pics.psych.stir.ac.uk/zips/ nottingham_originals.zip
24. Lyons, M.J., Akamatsu, S., Kamachi, M., Gyoba, J.: Coding Facial Expressions with Gabor Wavelets Proceedings, Third IEEE International Conference on Automatic Face and Gesture Recognition, April 14-16, Nara Japan, IEEE Computer Society (1998) 200–205
25. Georgia Tech face database. http://www.anefian.com/ research/face_reco.htm
26. Abeerden. http://pics.psych.stir.ac.uk/zips/Aberdeen.zip
27. Caltech Faces 1999. http://www.vision.caltech.edu/Image_Datasets/faces/faces.tar
28. Pain expression. http://pics.psych.stir.ac.uk/zips/pain.zip
29. The FEI face database. http://fei.edu.br/ cet/facedatabase.html
30. Forczmański, P.: Comparison of Tensor Unfolding Variants for 2DPCA-Based Color Facial Portraits Recognition. International Conference on Computer Vision and Graphics (ICCVG). LNCS 7594 (2012) 345–353