Pattern Recognition 46 (2013) 2634–2646

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

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

Local descriptors in application to the aging problem in face recognition

Michał Bereta [a,b], Paweł Karczmarek [a,c,*], Witold Pedrycz [a,d], Marek Reformat [a]

[a] Department of Electrical and Computer Engineering, University of Alberta, 9107-116 Street, Edmonton, AB, Canada T6G 2V4
[b] Institute of Computer Science, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland
[c] Institute of Mathematics and Computer Science, The John Paul II Catholic University of Lublin, Al. Racławickie 14, 20-950 Lublin, Poland
[d] Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland

ARTICLE INFO

Article history:
Received 8 June 2012
Received in revised form
16 January 2013
Accepted 13 March 2013
Available online 29 March 2013

Keywords:
Face recognition
Face detection
Aging process
Local descriptors
Local binary patterns
Gabor wavelets

ABSTRACT

Local descriptors are widely used in face recognition due to their robustness to changes in expression or occlusion in facial images. In this paper, a comparison of local descriptors commonly used in face recognition methods is presented in the context of age changes of individuals. We quantify abilities of local descriptors used in face recognition in the context of age discrimination. The performance of the descriptors is evaluated by experimenting with the FG-NET database. We present the results for different age groups and for various age differences of individuals present in the training and testing images. The values of recognition accuracy are reported in combination with various similarity measures used for classification purposes. Moreover, the performance of the descriptors combined with Gabor wavelet images is tested.

1. Introduction

Face recognition has received significant attention during the last years mainly because of its numerous applications in forensic science, border control, driver's license and passport verification, access control, localization of missing people, etc. Surveys of the face recognition methods can be found in [1,2].

"Traditional" face recognition studies focus on several factors such as influence of illumination, expression and pose on the performance of the classifiers. Age variation and age factor, in general are not commonly considered. One of the main reasons for a small number of studies concerning face recognition realized in the context of age factor was an absence of representative public databases with images containing individuals of different age as well as low quality of old photos. Two pertinent databases, namely MORPH and FG-NET (Face and Gesture Aging Database) have been recently published (cf. [3,4]) triggering studies in face recognition involving age factor.

Age-invariant face recognition methods comprise two main groups. The first of them embraces methods based on utilizing facial features being relatively robust to age progression. An example here is the method using gradient orientation pyramid (GOP) combined with a support vector machine (SVM) [5]. This algorithm was tested using two private passport datasets. The error rates are presented for different age gaps between the training and testing set. In [6], the study of the performance of principal component analysis (PCA) [7] and elastic bunch graph matching (EBGM) [8] is presented. The methods were tested separately and in combination, or being complemented by soft biometrics (e.g., race, gender.) The results were obtained for a large database containing individuals exhibiting significant age spans. It has been reported that the recognition accuracy does not degrade linearly with respect to age intervals. When the age differences are greater than 15 years, the accuracy becomes significantly lower than the accuracy reported for age differences within the period of 15 years. The study presented in [9] offers a comparison of PCA, GOP and Gabor wavelets followed by PCA, and local binary patterns (LBP) [10]. Here, the standard Gabor wavelets using five scales and eight orientations are shown as the best among the tested age invariant methods. LBP shows its good performance for age gaps between training and testing set in the range of 7–9 and 10–12 years. The results were obtained with the use of the Tchebyshev distance. Other analyses of PCA performance completed in the context of age progression can be found in [11,12]. In [13], a multi-feature discriminant analysis (MFDA) method was adopted to refine the feature space. An interesting concept of utilizing the most age-invariant facial region, i.e., periocular area,

* Corresponding author at: The John Paul II Catholic University of Lublin, Institute of Mathematics and Computer Science, Al. Racławickie 14, 20-950 Lublin, Poland. Tel.: +48 81 445 45 52; fax: +48 81 445 45 51.
E-mail addresses: beretam@gmail.com (M. Bereta), pawelk@kul.pl (P. Karczmarek), wpedrycz@ualberta.ca (W. Pedrycz), reformat@ualberta.ca (M. Reformat).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.03.010

was presented in [14], where the fusion of Walsh–Hadamard transform and LBP (so called Walsh–Hadamard Transform Encoded LBP, WLBP) was applied to this facial segment. Next, the unsupervised discriminant projection (UDP) [15] was used to build subspaces of the images.

The methods belonging to the second class of the methods are focused on age estimation and a simulation of the aging process itself [16–20]. These methods produce new images of the compared faces by using the aging model to compensate for the age effect. Next, such virtual faces are matched. Designing an aging model is still a challenging problem because of the fact that two people may age in a different way. Aging is affected by many uncontrolled factors (climate, lifestyle, living place, health, etc.) Moreover, aging models are difficult to implement.

The main objective of this study is to assess and compare the accuracy of the most recent local descriptors along with their application to age-invariant face recognition. The local approach seems to be one of the leading trends in facial recognition during the recent years, we can refer here e.g., to local binary pattern and its numerous modifications. The well-known advantage of the local descriptors is that they are relatively robust to changes in expression, luminance, and pose. We examine their robustness to age changes. It is worth noting that not only the LBP descriptor is studied here but we analyze the broader class of local descriptors including Gabor filters. In this way, the results of this study offer the through insight into the role and effectiveness of local descriptors in the context of age invariant face recognition. Using the FG-NET database, we are interested in evaluating the dependence of recognition rate on the age groups and age differences between the testing and training set. In particular, an estimation of the dependence of the recognition accuracy vis-à-vis the age of the subject comes as an interesting topic that could shed light on a potential applicability of local descriptors as functional components of more sophisticated facial recognition models exploiting other techniques such as subspaces, and wavelets. Furthermore, local descriptors could play an important role in age estimation, and in modeling of the aging processes. A detailed evaluation of the performance of local descriptors can be sought as a useful contribution of this study. Moreover, the paper contributes to the body of knowledge on age-invariant facial recognition by bringing forward a comprehensive, in-depth analysis of local descriptors.

In this research, we have tested the following commonly encountered descriptors: local binary pattern (LBP), improved LBP (ILBP), center-symmetric LBP (CSLBP), differential local ternary pattern (DLTP), three patch LBP (TPLBP), multi-scale block LBP (MBLBP), simplified form of Weber local descriptor (WLD), local XOR pattern (LXP), local Gabor phase binary pattern (LGPBP), local Gabor XOR pattern (LGXP). Moreover, the first seven of above descriptors were tested in application to Gabor magnitude images. This selection is motivated by the fact that the abovementioned descriptors come with the most promising accuracy in comparison with the accuracy obtained by the other local descriptors when using images of individuals without any age gap.

In this study, we discuss only the local descriptors and do not compare them with other approaches, e.g., combined methods. Such comparative analysis, albeit interesting per se, goes beyond the scope of this study. We focus on the aging aspect and its influence on the performance of local descriptors. Quantifying this dependence we are interested in identifying the best local descriptor, which can be potentially combined with other techniques, e.g., a global approach.

To classify an image one needs to compare it with the known images by calculating a distance between feature vectors containing local descriptions of these images. A sound choice of the distance used for this purpose can improve recognition rate. Therefore, to determine a suitable distance, we compare recognition accuracies for the best local descriptors involving various similarity/dissimilarity measures such as Euclidean, cosine, correlation, Manhattan, Tchebyshev, Bray–Curtis, and Canberra distances and also more specialized measures, namely histogram intersection, chi square statistics, and log-likelihood statistics. The nearest neighbor classifier is used in the recognition phase.

The paper is organized as follows. Local descriptors under study are presented in Section 2. Here we describe Gabor wavelets and their combinations with local descriptions of images. Moreover, similarity/dissimilarity measures used to calculate distances between feature vectors are presented. In Section 3, we cover a description of aging database FG-NET, which is commonly used in facial aging considerations. Section 4 contains results of experiments while conclusions are covered in Section 5.

2. Local descriptors for facial recognition

2.1. Local descriptors

For many years local descriptors have been developed in texture analysis. Just recently many of them have been investigated and applied to facial recognition. The idea of local description can be outlined as follows. Local features of the image are described using neighboring pixels. Next, they are aggregated to form a general description of the entier image. In this way, we form a vector of features to be compared with a vector of an unknown image. Sometimes this vector is constructed with features histograms for the blocks, which the image is divided into. These histograms are constructed from pixels' labels (i.e., descriptions.) For a given pixel, its label is calculated using its neighborhood. A concatenation of all block descriptions gives rise to the global geometry of face encoded in a single vector. The advantages of such approach come with its robustness to changes in expression, luminance or pose. Furthermore the approach does not require a training stage. Its shortcoming may be the size of features vectors, which could make the recognition process computationally expensive. Fig. 1 illustrates a general concept of the local description. The ith fiducial point ($i = 1, \ldots, n$) of an image is described by the vector (particularly it can be a single number) v_i. Next, all such vectors are concatenated into a single image description v.

Let us briefly characterize the local descriptors considered in this paper. One of the best known local descriptors is a local binary pattern (LBP), presented in face recognition context in [10]. The generic formula for determining a local binary description for a given pixel p_c comes in the following form:

$$\mathrm{LBP}(p_c) = \sum_{i=0}^{7} s(p_i - p_c)2^i, \quad \text{where} \quad s(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases}$$

In what follows p_c and p_i, $i = 0, \ldots, 7$, are the grey-level values of the central pixel and its neighbors, respectively. An example of basic LBP operation for a given pixel is depicted in Fig. 2. A threshold operation $s(\cdot)$ transforms elements of the matrix constructed from the values of the central pixel and its neighbors into binary numbers by comparing them with the central element's value. The numbers obtained in this manner are then concatenated in a clockwise direction and a final label of the central pixel is calculated. Fig. 3 presents the result of LBP operation for each pixel of an example image coming from the FGNET database.

The LBP operator is often used in conjunction with a circular neighborhood [10]. In this case the values of pixels are usually bilinearly interpolated at sampling equidistant points on a circle with a center at the point p_c. The LBP descriptor is robust to illumination changes and relatively easy to compute.