Pattern Recognition Letters 36 (2014) 177–185

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Class specific subspace dependent nonlinear correlation filtering for illumination tolerant face recognition

Pradipta K. Banerjee [a], Asit K. Datta [b,*]

[a] Department of Electrical Engineering, Future Institute of Engineering and Management, Kolkata 700150, India
[b] Department of Applied Optics and Photonics, University of Calcutta, 92 Acharya Prafulla Chandra Road, Kolkata 700009, India

ARTICLE INFO

Article history:
Received 17 February 2013
Available online 24 October 2013

Communicated by M.S. Nixon

Keywords:
Correlation filter
Face recognition
Illumination invariant
Class-specific subspace

ABSTRACT

A frequency domain nonlinear correlation technique for face recognition under varying lighting conditions is proposed. The technique is based on phase correlation between an optimum projecting image correlation filter and an optimum reconstructed image correlation filter during class specific subspace operation. Performance improvement is achieved by exploiting point wise nonlinearities of image pixels. Further optimization is carried out by minimizing the energy at the correlation plane while maximizing the correlation peak. While comparing with other standard unconstrained correlation filters, improved performance of the proposed scheme is established by experimental results on standard face data bases.

1. Introduction

Several linear correlation filters and their variants were developed over the last few years (Mahalanobis et al., 1987; Mahalanobis et al., 1994; Refregier, 1991). These are not only successfully applied for pattern recognition but also for used for face recognition tasks (Kumar et al., 2004; Lai et al., 2008; Rizo-Rodrguez et al., 2012; Banerjee and Datta, 2013). In some cases, to achieve better performance under varying illumination conditions during face recognition, simple preprocessing in spatial domain (del Solar and Quinteros, 2008), logarithmic pre-processing in frequency domain (Savvides and Kumar, 2003), Retinex pre-processing (Levine and Yu, 2006) and training based pre-processing (Banerjee and Datta, 2013) were carried out before correlation filtering operations. It has been noted, that in most of the cases of illumination invariant face recognition (Maddah and Mozaffari, 2012; Jeong et al., 2009; Levine and Yu, 2006), the maximum average correlation energy (MACE) filters or unconstrained MACE (UMACE) filters (Mahalanobis et al., 1994) were used. These filters were designed with a set of randomly or systematically chosen training images, so that the designed filter could exhibit precise classification while unknown illumination variations are present in test faces.

It is however, not always possible to select a set of proper training images so that all conditions of illumination variation in test face image will lie in the convex hull of training variations (Banerjee and Datta, 2013). Increasing the number of training images can provide a solution, though it has been reported (Kumar and Pochapsky, 1986) that the signal to noise ratio (SNR) of such a filter monotonically decreases with the increase in the number of training images. A solution to this problem may be addressed, if the nature of the correlation filter is changed dynamically according to the input test face images so as to achieve robust recognition for all possible illumination variations that lie in a three dimensional (3D) linear subspace for a Lambertian model. Towards achieving this goal, the proposed method takes the help of face reconstruction using class specific subspace analysis. It has been shown in Liu et al. (2003) that the low energy in the residue image can be a good criterion to authenticate a face image and thus it is possible to better performance if class specific subspace analysis is performed instead of classical subspace analysis.

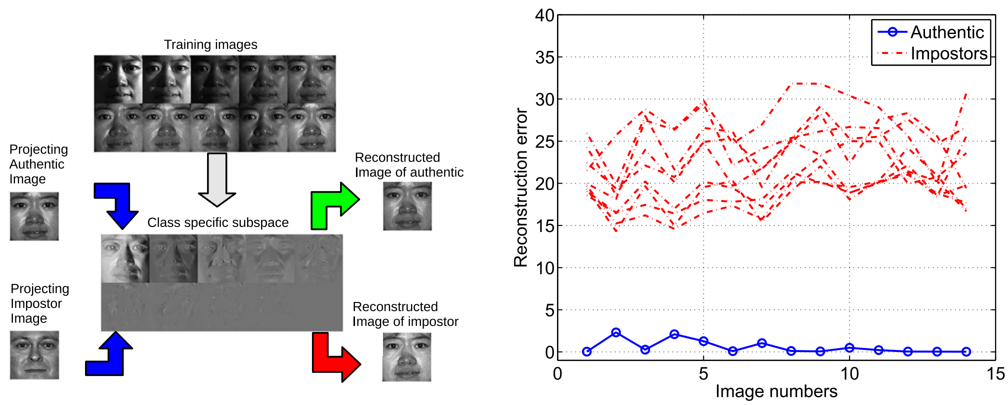
Fig. 1(a) shows a basic structure of face reconstruction process. It has been shown, that the test face image is almost perfectly reconstructed when the test face is taken from a class of training faces. It is interesting to note from Fig. 1(a) that, the image of person-3 (taken from PIE database Sim et al., 2003) when projected onto the subspace developed by person-1 images, the reconstructed image looks like person-1 after reconstruction. That is, the orthonormal eigenface basis of each individual best spans the face of the same person rather than the face of other persons. This statement can be justified by computing the reconstruction error defines as the squared norm between test face and its reconstructed version and is given in Eq. (1) while projected on some one's (say person-1 here) subspace.

$$Error = \|t - r\|^2 \tag{1}$$

where t is test face and r is its reconstruction.

* Corresponding author.
E-mail addresses: pradiptak.banerjee@gmail.com (P.K. Banerjee), asitdatta@gmail.com (A.K. Datta).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.10.012



(a) Face reconstruction (b) Error plot for all persons

Fig. 1. (a) The face reconstruction is shown while the subspace analysis is performed over an individual (person-1) from PIE face. (b) Reconstruction error plot for person-1's individual eigenface subspace. Large errors found for impostors and minimum errors for authentic (person-1).

If the error is below a certain specified value, all faces from person-1 can be treated as authentic and the faces of other persons, except person-1, are treated as impostors. Fig. 1(b) shows the reconstruction errors obtained for both authentic and impostor faces. From Fig. 1(b) it is observed that large reconstruction errors are obtained for impostors compared to authentic face images. Therefore, authentic face images and their variations are expected to be modeled well in their individual eigenface subspace, thus leading to small reconstruction errors and similarly impostor face images are not well represented by some one else's individual eigenface subspace and should result in larger reconstruction error. It is also observed from Fig. 1(a) the impostor reconstruction is inferior than the authentic reconstruction. Therefore, a reconstruction which helps in discriminating an authentic or an impostor face must be a component for an improved filter design. The phase correlation of filters synthesized from projecting image and reconstructed image gives sharp peak for authentic and no such peak for impostor.

There are other problem of using linear correlation filters for face recognition purpose. The linear filter formulation considers images having nonuniform dynamic range and hence in the testing stage, it is hard to discriminate authentic and impostor images that lie below a span of low gray level. To overcome this situation this paper proposes nonlinear correlation filter by exploiting the point nonlinearites (Mahalanobis and Kumar, 1997) of image pixels so that the designed correlation filter achieves a uniform dynamic range. This type of nonlinear mapping stretches pixel distribution of face images in a wide range and consequently high frequency components are amplified.

In this study three approaches are judicially combined to improve face recognition results under illumination variation viz, (i) projection based method of designing correlation filter is used to improve upon the capability of recognition at all possible illumination variations (ii) phase correlation method is used to enhance peak sharpness at the correlation plane for authentic face image as the phase contains more information than the magnitude of Fourier spectrum and (iii) point nonlinearities are considered to extend uniform dynamic range. Two correlation filters are designed, in order to achieve this amalgamation, i.e. (a) nonlinear optimum projecting image correlation filter $\mathbf{H}_p$, and (b) nonlinear optimum reconstructed image correlation filter $\mathbf{H}_r$. The nature of design process of these two filters is same with only difference in the images are used for synthesis. Design of $\mathbf{H}_p$ uses projecting image and the design of $\mathbf{H}_r$ includes reconstructed image. The phase correlation between these two filters produces a response surface. The nature of the surface totally depends on the face class involved. Ideally, a delta type peak at the correlation plane is obtained if these two filters are generated from the same face class. Experimental results on standard databases like YaleB (extended) (Lee et al., 2005) and PIE database (Sim et al., 2003) witness the promising performance of the proposed method when compared to other standard correlation filter based face recognition systems.

2. Formulation of nonlinear correlation filters

2.1. Nonlinear optimum projecting image correlation filter (NOPICF)

Any projecting image from kth class (where, $k = 1, 2, \ldots, M$) is represented in spatial domain as T (in matrix form of size $d_1 \times d_2$) or t (in vector form of size $d \times 1$, where $d = d_1 \times d_2$) and its frequency domain counterparts are $\mathbf{T}$ and $\mathbf{t}$, respectively. $\bar{\mathbf{T}}$ represents the diagonal form of $\mathbf{t}$. The point wise nonlinearities of an image can be achieved according to power law transformation given by,

$$t_\alpha^\beta = \alpha[t]^\beta \tag{2}$$

where, $\alpha > 0$ and $\beta > 0$, can be integer or fraction.

Eq. (2) indicates that each element of t is scaled by αth amount and raised to βth power. Hence for $\alpha = \beta = 1$, the image t_1^1 represents the original image t. If $\mathbf{h}_{\alpha\beta}$ is an optimum correlation filter corresponding to projecting image $\mathbf{t}_\alpha^\beta$, then the correlation plane $\mathbf{g}_{\alpha\beta}$ in response to $\mathbf{t}_\alpha^\beta$ is given by,

$$\mathbf{g}_{\alpha\beta} = \bar{\mathbf{T}}_\alpha^{\beta*}\mathbf{h}_{\alpha\beta} \tag{3}$$

where * represents conjugation operation.

From Eq. (3), it may be noted that a number of correlation planes $\mathbf{g}_{\alpha\beta}$ as well as a number of classifiers $\mathbf{h}_{\alpha\beta}$ is generated for each value of $\alpha = \alpha_1, \ldots, \alpha_n$ and $\beta = \beta_1, \ldots, \beta_m$ in response to a single projecting image. The same can be denoted as, $\bar{\mathbf{T}}_{\alpha_1}^{\beta_1}, \bar{\mathbf{T}}_{\alpha_1}^{\beta_2} \ldots, \bar{\mathbf{T}}_{\alpha_1}^{\beta_m}, \bar{\mathbf{T}}_{\alpha_2}^{\beta_1}, \bar{\mathbf{T}}_{\alpha_2}^{\beta_2}, \ldots, \bar{\mathbf{T}}_{\alpha_2}^{\beta_m}, \ldots, \bar{\mathbf{T}}_{\alpha_n}^{\beta_1}, \bar{\mathbf{T}}_{\alpha_n}^{\beta_2}, \ldots, \bar{\mathbf{T}}_{\alpha_n}^{\beta_m}$. Hence from Eq. (3), a set of correlation planes can be derived as,

$$\begin{aligned}
\mathbf{g}_{\alpha_1\beta_1} &= \bar{\mathbf{T}}_{\alpha_1}^{\beta_1*}\mathbf{h}_{\alpha_1\beta_1} \\
\mathbf{g}_{\alpha_1\beta_2} &= \bar{\mathbf{T}}_{\alpha_1}^{\beta_2*}\mathbf{h}_{\alpha_1\beta_2} \\
\vdots &= \vdots \\
\mathbf{g}_{\alpha_n\beta_m} &= \bar{\mathbf{T}}_{\alpha_n}^{\beta_m*}\mathbf{h}_{\alpha_n\beta_m}
\end{aligned} \tag{4}$$

Since a sharp and distinct correlation peak in correlation plane reduces the chances of misclassification, minimization of energy at the correlation plane containing undesired side lobes and