COLOR HOG-EBGM FOR FACE RECOGNITION

David Monzo   Alberto Albiol Antonio Albiol Jose M. Mossi

ITEAM Institute, Universidad Politécnica de Valencia (Spain)

ABSTRACT

Most face recognition algorithms make only use of intensity information of the images discarding color as a distinctive cue. This paper extends the HOG-EBGM face recognition algorithm to use color information. In HOG-EBGM, each face is represented by a graph described by HOG features at specific landmarks. The color extension of the method here proposed intends to make the algorithm robust against color changes, while keeping its former robustness against scale, position, rotation and intensity variations. In the paper, several color representations of the faces are studied. Also, to reduce the higher dimensionality of the new color features, a comparison of dimensionality reduction techniques is included. Results on the Experiment 4 of the Face Recognition Grand Challenge show that color HOG-EBGM outperforms the gray-scale version of the algorithm in all cases. The best results were obtained using the Opponent color space with LDA.

1. INTRODUCTION

Face recognition is a task that has attracted much attention among researchers during the last years. Although a large number of approaches have been proposed (we recommend [14] for an extensive survey on the topic), the majority of them have only explored the use of intensity-based information to represent faces. However, color is generally considered a powerful cue for object recognition and it has been proved that it can be a distinctive feature between different faces [1]. While illumination variance can affect the recognition performance, the extraction of color descriptors can help to achieve invariance to the changes of illumination.

Authors have recently worked on the study of efficient color models which help to overcome the color invariance problem [5], particularly, they have focused on getting color spaces suitable for the extraction of discriminant color features. In [13] the authors explore the most conventional color spaces to finally probe that for each model descriptors achieve different grades of invariance to light intensity and light color changes. In [7] there is an analysis of the optimal way of combining the color components of images to get monochromatic images minimizing the information loss for face recognition. In that line, Yang and Liu [17] propose a discriminant criterion to create a color space in which the color components are extracted to best represent images for recognition purposes by combination of the RGB components.

Many successful face recognition algorithms originally developed for gray-scale images have been afterwards extended to use color information. This is the case of PCA and LDA, two of the most well-known algorithms for face recognition. Some examples of color PCA and LDA approaches can be found in [12, 15]. Also, some powerful object descriptors such as SIFT[8] or HOG [9] which initially were developed for intensity images have also been extended to embed color information [2, 13].

Nevertheless, still there is a number of face recognition algorithms which have been reported to have good performance in gray-scale and have never been studied to work in a color space. This is the case of the HOG-EBGM algorithm, a novel face recognition method recently proposed in [3]. HOG-EBGM is a variation of Wiskott's Gabor-EBGM method [16], which has become a standard baseline algorithm for comparison of face recognition algorithms. EBGM methods describe each face by the information extracted from the nodes of a face graph (FG) which is automatically placed on facial landmarks. The main difference between HOG-EBGM and Gabor-EBGM is the nature of the local features used in each node graph: HOG descriptors and Gabor jets respectively. HOG descriptors are based on image gradients in local areas, while Gabor jets are built using Gabor wavelet coefficients. HOG descriptors have been proved to be robust against small displacements, rotations and intensity and non-linear illumination changes around facial landmarks, making them suitable for recognition tasks. Interested readers can see [10] for a performance comparison between both algorithms.

The motivation of this paper is to extend our HOG-EBGM algorithm to use color information. This adaption aims to give the algorithm more robustness and discrimination power against color changes, while keeping its former robustness against scale, position, rotation and light variations. Also, in this work we carry out a study that explores different color models in order to find the optimum color space for the recognition task.

The rest of the paper is organized as follows. In Section 2 the original HOG-EBGM is described. Next, Section 3

This work was funded by the Spanish Government CYCIT grant TEC2009-09146.

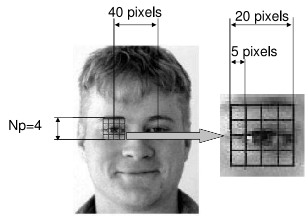

Fig. 1. Normalized face and the spatial bins of the right eye HOG descriptor.

extends the algorithm to use color, discussing about different color representations and dimensionality reduction techniques. Finally, results and conclusions are presented in Sections 4 and 5.

2. GRAY-SCALE HOG-EBGM

In HOG-EBGM (see [3] for implementation details) facial landmarks are located on a graph that follows the structure proposed in the CSU project [4]. The number of nodes in this graph is 25. Each landmark is modeled using a set of training faces that account changes of expression, hair styles, illumination, etc. These models are used to automatically locate landmarks on unseen faces. This automatic detection is done through an iterative process in which the position of a new landmark is predicted using previously detected landmarks to reduce the search area. The process to detect the i-th ($i > 2$) facial landmark (summarized in a few lines due to space restrictions) is:

1. Coarse initial prediction of the facial landmark location X_i^s, based on the geometry of the previously detected $j < i$ points in the FG.

2. Refine the previous prediction on a reduced search area using the landmark model.

Each HOG descriptor is a histogram in which the bins form a three dimensional lattice with $N_p = 4$ bins for each spatial direction and $N_o = 8$ bins for the orientation for a total of $N_p^2 N_o = 128$ components. In our work, each spatial bin is a 5×5 pixels square (Fig. 1). The contribution of each pixel gradient to the histogram is weighted by a Gaussian window centered at the keypoint and trilinearly interpolated with the surrounding bins. See [3] for minor implementation details.

After the location of the graph nodes, each landmark is described by:

$$J_i = \text{HOG}(I, X_i) \in R^{N_{hog}}, N_{hog} = 128 \quad (1)$$

where I is the intensity face image. For recognition tasks, a face F is finally represented by the vector resulting of the concatenation of the descriptors associated to the 25 landmarks of the FG:

$$F = [J_1, J_2, \ldots, J_{25}] \in R^{N_f} \quad (2)$$

where $N_f = N_{hog} \times 25 = 3200$.

3. COLOR HOG-EBGM

In this section we extend the HOG-EBGM algorithm to use color information. The resulting approach will be called Color HOG-EBGM or simply CHOG-EBGM hereafter. The fundamental change between gray-scale and color HOG-EBGM algorithms is that the descriptors J_i are replaced by new J_i^c descriptors:

$$J_i^c = [HOG(C_1, X_i), HOG(C_2, X_i), HOG(C_3, X_i)]$$

where C_1, C_2 and C_3 are each of the color channels in a suitable color space. Section 3.1 will discuss some of the color spaces used in this paper. Notice that $J_i^c \in R^{N_{chog}}$, where $N_{chog} = N_{hog} \times 3 = 384$.

For recognition tasks, a color face F_C is represented by the vector resulting of the concatenation of the J_i^c,$1 < i < 25$ descriptors:

$$F_C = [J_1^c, J_2^c, \ldots, J_{25}^c] \in R^{N_c} \quad (3)$$

with $N_c = N_{chog} \times 25 = 9600$. As in the case of HOG-EBGM the dimensionality of F_C is quite high and dimensionality reduction is required. Section 3.2 will review the techniques used to this end.

3.1. Color Models

One major issue when considering color for the extraction of features is the actual color model used to represent the image, described through its channels C_1, C_2 and C_3. Many models exist, but some authors [13, 17] have focused on those that provide more discriminant information when using descriptors for the separability of different classes. As it is not an easy task to determine which model fits better with the requirements of face recognition, in this work we have experimented with three conventional models used in other works for recognition tasks (RGB, HSV and Opponent Color) and we have also compared them with a particular model proposed by Yang and Liu in [17]. We will only describe here the Opponent Color and Discriminant Color spaces:

Opponent Color model This model, derived from the RGB channels, and is defined by:

$$(C_1, C_2, C_3)^T = \left(\frac{R-G}{\sqrt{2}}, \frac{R+G-2B}{\sqrt{6}}, \frac{R+G+B}{\sqrt{3}} \right)^T \quad (4)$$

where C_1 and C_2 have color information and are shift invariant to light intensity, while C_3 represents the intensity.

Discriminant Color Space (DCS) [17] This is a color model specifically developed for recognition purposes. The idea behind this model is that a suitable color space can be learned from a set of model by optimally combining the RGB

components. For the computation of the optimal weights, this model uses the Fisher criterion which maximizes the between-scatter matrix and minimizes the within-scatter matrix using the three-dimensional color pixels as input vectors for a particular dataset.

3.2. Reduction of dimensionality

Due to the high dimension of the face vectors extracted with HOG-EBGM and CHOG-EBGM, we apply dimensionality reduction prior to recognition. In this paper, we have studied three algorithms: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) [6] and Orthogonal Linear Discriminant Analysis (OLDA) [18]. In this paper, we will only explain the OLDA method which is seldom known.

OLDA is one of the many variations of LDA to deal with the undersampling problem. The key idea of OLDA is that the discriminant vectors are orthogonal to each other. In [18] Ye provides an efficient way of computing OLDA. The algorithm starts by obtaining the non-null space from the total scatter matrix S_t:

$$S_t = [U_1, U_2] \begin{pmatrix} \Sigma_t^2 & 0 \\ 0 & 0 \end{pmatrix} [U_1, U_2]^T \tag{5}$$

where U_1 spans the non-null space of S_t and Σ_t^2 are the eigenvalues corresponding to this subspace. Next, the class centered data matrix H_b is projected onto the U_1 subspace:

$$B = \Sigma_t^{-1} U_1^t H_b \tag{6}$$

Then, SVD decomposition is applied to B:

$$B = P \hat{\Sigma} Q^T \tag{7}$$

This allows to obtain a projection matrix X that simultaneously diagonalizes S_b, S_w and S_t:

$$X = U \begin{pmatrix} \Sigma_t^{-1} P & 0 \\ 0 & I \end{pmatrix} \tag{8}$$

Since, usually $\mathrm{rank}(S_b) = C - 1 = q$, the next step takes the first q columns of X to obtain a new projection matrix X_q. In [18] Ye demonstrated that the Fisher criterion is maximized with $\Phi = X_q M$, where M can be any arbitrary $q \times q$ nonsingular matrix. This degree of freedom to choose M allows to create an orthogonal matrix Φ by setting $M = \hat{R}$, where $\hat{R}$ comes from the QR decomposition of X_q.

4. EXPERIMENTS AND RESULTS

The performance of the proposed algorithm is evaluated in this paper using the Face Recognition Grand Challenge (FRGC) [11] database, which is one of the most challenging and large-scale face databases. The FRGC evaluation protocol is organized in a set of experiments. We choose the Experiment 4 which is designed for comparing face recognition

Training **Target** **Query**

Fig. 2. Normalized face images of the FRGCv2 Exp. 4.

algorithms under uncontrolled acquisition conditions. This is by far the most challenging experiment in FRGCv2 as shown by some competition results [11]. More in detail, FRGCv2 Experiment 4 consists of 12,766 training images, 16,028 controlled target images and 8,014 uncontrolled query images. Due to the large quantity of training color images, we approximately halved the training set with a random selection of training samples. This allows us to reduce the computational burden and memory allocation problems. Fig. 2 shows and example of normalized training, target and query of the FRGCv2 Experiment 4.

The gray-scale HOG-EBGM and the CHOG-EBGM are compared using two different methods: for recognition we use rank curves, while for verification we use a Receiver Operating Characteristic (ROC) curve proposed in the FRGC protocol. Also, to study the influence of the color space representation we obtained the previous curves for each color model: RGB, HSV, Opponent Color and DCS. From [17], DCS color space conversion from RGB in FRGCv2 is:

$$X_{FRGC} = \begin{pmatrix} 0.28 & 0.06 & 0.74 \\ 0.06 & -0.16 & 0.40 \\ 0.29 & -0.21 & -0.12 \end{pmatrix}$$

To measure similarity between projected face vectors, we can use many different similarity measurements. In this paper, only results using cosine distance are presented here since they clearly outperform all results obtained using euclidean distance.

Figure 3 shows the rank curves resulting for the different approaches. The results of the ROC curves for verification regarding a false accept rate of 0.1% are shown in Table 1. The rank and the ROC curves show a higher performance of CHOG-EBGM when compared to gray-scale HOG-EBGM, regardless of the color space in use. Also, in all the experiments the Opponent Color space outperforms the results obtained with the rest of the color models.

Regarding the dimensionality reduction applied to the data for the recognition and verification tasks, both LDA algorithms have proved to be much more suitable than PCA for matching the descriptor vectors of the face graphs. Let's also notice that for recognition, OLDA achieves the highest rates

	PCA	LDA	OLDA
Gray-scale HOG-EBGM	7.7%	53.9%	37.0%
CHOG-EBGM on RGB	7.8%	56.8%	42.6%
CHOG-EBGM on HSV	9.5%	55.5%	44.0%
CHOG-EBGM on Opp. Color	10.1%	57.4%	44.4%
CHOG-EBGM on DCS [17]	9.5%	54.5%	40.9%

Table 1. FRGCv2 Exp. 4 Verif. Rate at 0.1% of FAR.

while for verification the best rates are obtained with the baseline LDA.

5. CONCLUSIONS

In this work we present a new method, Color HOG-EBGM which is derived from the gray-scale HOG-EBGM. Experiments show that better performance is achieved in the color version for all the color spaces studied. Also, from all these color models, Opposite Color is the most effective, as it achieves the highest recognition rates, even outperforming specific methods designed for recognition tasks, as the one proposed in [17].

6. REFERENCES

[1] P. S. A. Yip. Role of color in face recognition. *Journal of Vision*, 2(7):596–596, November 2002.

[2] A. E. Abdel-Hakim and A. A. Farag. Csift: A sift descriptor with color invariant characteristics. In *CVPR*, 2006.

[3] A. Albiol, D. Monzo, et al. Face recognition using hog-ebgm. *Pat. Rec. Letters*, 29(10):1537–1543, 2008.

[4] D. Bolme et al. The CSU face identification evaluation system In *Comp. Vision Systems*, Apr. 2003.

[5] D. H. Brainard and W. T. Freeman. Bayesian color constancy. *Journal of Optical Society of America*, 14(7):1393–1411, 1997.

[6] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 1990.

[7] C. F. J. III and A. L. Abbott. Optimization of color conversion for face recognition. *EURASIP*, 2004:522–529, 2004.

[8] D. Lowe. Distinctive image features from scale-invariant keypoints. *Journal of Comp. Vision*, 60(2):91–110, Nov. 2004.

[9] M. Bicego. et al. On the use of SIFT features for face authentication. In *CVPRW*, June 2006.

[10] D. Monzo, A. Albiol, and J. Sastre. Hog-ebgm vs. gabor-ebgm. In *ICIP*, pages 1636–1639, October 2008.

[11] P. J. Phillips, et al. Overview of the face recognition grand challenge. In *CVPR*, 1:947–954, June 2005.

[12] M. Thomas, et al. Face recognition using a color pca framework. In *Comp. Vision Systems*, pages 372–382, 2008.

[13] K. E. A. van de Sande, et al. Evaluation of color descriptors for object and scene recognition. In *CVPR*, 2008.

[14] Z. W. and et al. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, Dec. 2003.

[15] C. Wang, B. Yin, X. Bai, and Y. Sun. Color face recognition based on 2dpca. In *ICPR*, pages 1–4, 2008.

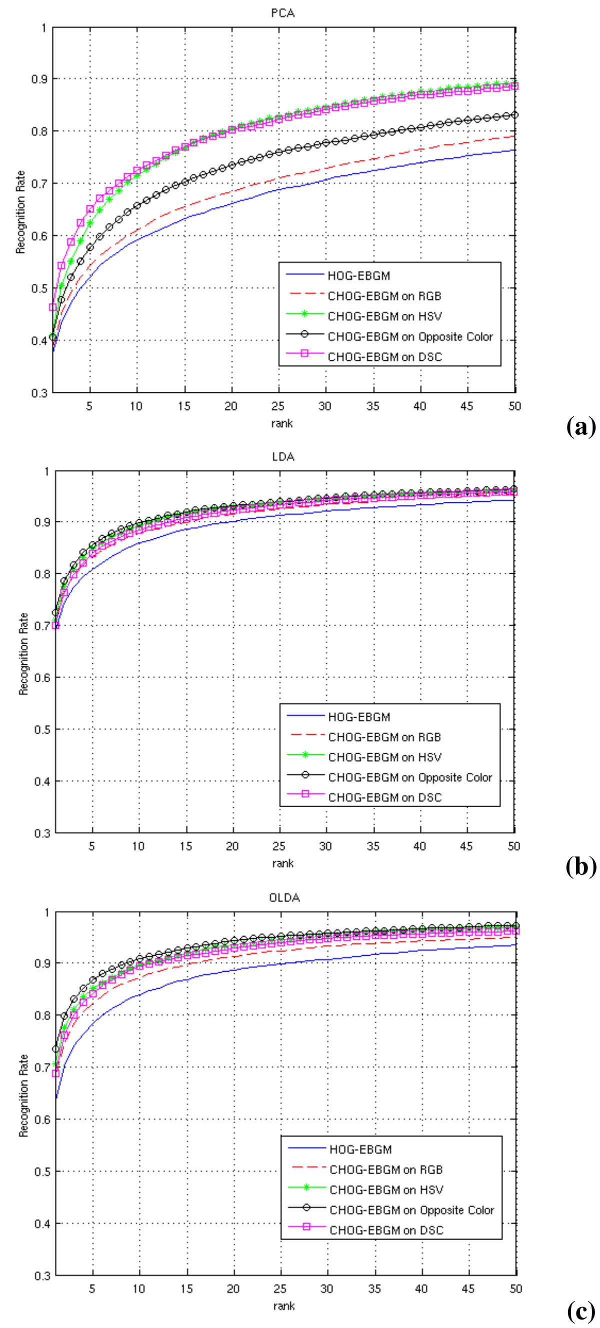

Fig. 3. Rank Curves for HOG-EGBM and CHOG-EBGM in different color spaces using a) PCA, b) LDA and c) OLDA.

[16] L. Wiskott, et al. Face recognition by EBGM. Tech. Report, Ruhr-Universitat Bochum, April 1996.

[17] J. Yang and C. Liu. A discriminant color space method for face representation and verification on a large-scale database. In *ICPR*, pages 1–4, 2008.

[18] J. Ye. Characterization of a family of algorithms for generalized discriminant analysis on undersampled problems. *Journal of Machine Learning Research*, 6:483–502, 2005.