

## Accepted Manuscript

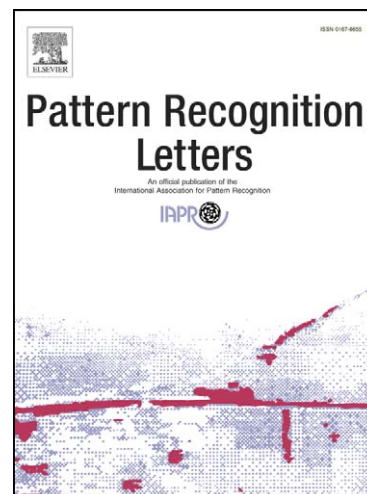
Face recognition using HOG-EBGM

Alberto Albiol, David Monzo, Antoine Martin, Jorge Sastre, Antonio Albiol

PII: S0167-8655(08)00110-4  
DOI: [10.1016/j.patrec.2008.03.017](https://doi.org/10.1016/j.patrec.2008.03.017)  
Reference: PATREC 4391

To appear in: *Pattern Recognition Letters*

Received Date: 25 May 2007  
Revised Date: 10 January 2008  
Accepted Date: 14 March 2008



Please cite this article as: Albiol, A., Monzo, D., Martin, A., Sastre, J., Albiol, A., Face recognition using HOG-EBGM, *Pattern Recognition Letters* (2008), doi: [10.1016/j.patrec.2008.03.017](https://doi.org/10.1016/j.patrec.2008.03.017)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Face recognition using HOG-EBGM

Alberto Albiol, David Monzo, Antoine Martin, Jorge Sastre,  
Antonio Albiol

*I-TEAM, Universidad Politecnica de Valencia, Spain*

---

## Abstract

This paper presents a new face recognition algorithm based on the well-known EBGM which replaces Gabor features by HOG descriptors. The recognition results show a better performance of our approach compared to other face recognition approaches using public available databases. This better performance is explained by the properties of HOG descriptors which are more robust to changes in illumination, rotation and small displacements, and to the higher accuracy of the face graphs obtained compared to classical Gabor-EBGM ones.

*Key words:* face recognition, EBGM, SIFT, HOG, local image descriptors

---

## 1 Introduction

Face recognition is a research field that has attracted much attention in the past years due to its many applications in public surveillance, video annotation, multimedia and others.

Many algorithms have been proposed for face recognition, we recommend (1) for a complete survey on the topic. Face recognition techniques can be broadly classified into holistic and feature based. Holistic methods, such as PCA (2) or LDA (3), project input faces onto a dimensional reduced space where recognition is carried out. These type of methods can be deemed as classical today. However, they are still very popular due to their simplicity and good performance. The greatest problem of eigenbased methods is that in some sense

---

\* This work has been supported by the Technical University of Valencia: Programa de apoyo a la investigación y desarrollo, PAID-06-06.

\* Corresponding author.

*Email address:* `alalbiol@dccom.upv.es` (Alberto Albiol).

they assume that faces are rigid objects which can be reconstructed with linear combinations of a set of eigenfaces or fisherfaces, which is not true. Also when the illumination conditions change the performance of these methods degrades rapidly. To overcome these difficulties many variations of the original methods have been proposed (4).

Feature based methods try to recognize faces using its facial components: eyes, nose, mouth, etc. Active Appearance Model (5) and Elastic Bunch Graph Matching (EBGM) (6) fall into this category. In EBGM faces are represented as graphs with nodes at facial landmarks (such as eyes, tip of the nose, etc.). Each node contains a set of Gabor wavelet coefficients, known as a jet. To increase robustness of EBGM to changes in expression and illumination, new approaches have been proposed. For example, in (7) a graph matching approach which replaces the original Gabor features is presented. Following this line of research, in this paper we present a EBGM algorithm in which Gabor features have been replaced by Histograms of Oriented Gradients (HOG) (8) descriptors, which are inherited from the Scale Invariant Feature Transform (SIFT) proposed by Lowe (9). SIFT have emerged as a cutting edge technology for extracting distinctive features from images, to be used in algorithms for tasks like matching different views of an object or scene. SIFT achieves invariance to scale changes by extracting keypoints at the local extrema of the scale space representation of the image, then each keypoint is represented using histograms of image gradients, in the sequel HOG descriptor. HOG descriptors have also been proposed for pedestrian detection (10; 11; 12). In these approaches objects are assumed to be at a fixed scale and are divided into small connected regions at fixed positions. Then, for each region a HOG descriptor is obtained and the combination of these descriptors is used to represent the object.

SIFT has also been recently proposed for face recognition (8), however this approach totally differs from ours. In (8) keypoints are located at the local extrema of the scale space as in the original Lowe's approach (9). The main problem of this approach is that there is no control on the number, position and scale of the keypoints. However, in our algorithm the keypoints represent specific facial landmarks which are detected first as explained below. Once facial landmarks are detected we use HOG descriptors to represent them.

The rest of the paper is organized as follows. First, Section 2 presents in detail how HOG descriptors are built. Next, Section 3 describes our EBGM algorithm that uses HOG descriptors. Sections 4 and 5 show the experimental setup and recognition results. Finally some conclusions and future research are drawn in Section 6.

Henceforth we will use the term Gabor-EBGM, or simply EBGM, when referring to the original EBGM algorithm presented in (6) while we will use the

term HOG-EBGM when referring to the algorithm presented in this paper.

## 2 HOG descriptors

As mentioned previously SIFT has emerged as one of the most used detection/description schemes for its ability to handle image transformations like scale changes (zoom), image rotation, and illumination. The major steps of the SIFT algorithm are:

- (1) Scale-space extrema detection
- (2) Orientation assignment
- (3) Keypoint descriptor

The first step is used by SIFT to achieve invariance to scale changes. This is done by extracting SIFT features only at the local extrema of the scale-space representation of the image. The next step aims to obtain image rotation invariance. To that end, at each extrema of the scale-space representation, SIFT finds the dominant orientation using image gradient information and then, all image gradients are made relative to this dominant direction.

While these two techniques have proved to be very useful for images that are arbitrarily scaled or rotated, the fact is that these normalization stages remove information which might be useful for recognition when images are not scaled or rotated. In this paper, we assume that the exact location of both eyes is known a priori. To detect the eyes precisely, we have developed an algorithm that uses a mixed approach of boosted classifiers (13) and again HOG descriptors. However this problem can be deemed as precise face localization and it is not treated here. Since the exact location of the eyes is used to normalize faces, we do not expect any changes in either scale or rotation. For this reason, we skip the two first steps of the SIFT algorithm and only adopt the last step from Lowe's approach, the keypoint descriptor. This keypoint descriptor is also called HOG in the literature.

The HOG descriptor is a local statistic of the orientations of the image gradients around a keypoint. More formally, each descriptor is a bundle of histograms composed of pixel orientations given by their gradients. The number of possible orientations (histogram bins) is referred to as  $N_o$ . Each histogram in the bundle describes a specific area around the keypoint. These areas correspond to the cells of a  $N_p \times N_p$  squared grid centered on the keypoint (see Fig. 1). The original paper (9) sets the parameters of the descriptor to  $N_p = 4$  cells for each spatial direction and  $N_o = 8$  bins for each histogram in the bundle resulting in a total of  $N_p^2 N_o = 128$  elements in a HOG descriptor.

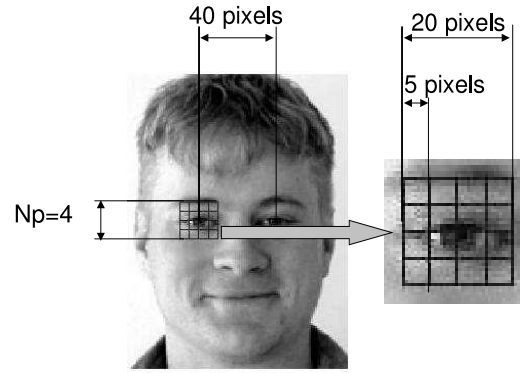


Fig. 1. Normalized face and the spatial cells of the right eye HOG descriptor

In our work, each spatial cell is a square of  $5 \times 5$  pixels. This size is chosen accordingly to the distance between eyes of the normalized faces, which in our work is 40 pixels. The results presented in Section 5.1.1 will further justify this selection.

Similar to Lowe's original approach, the contribution of each pixel gradient to the histogram is weighted by the gradient modulus and a Gaussian window. The Gaussian window is centered at the keypoint coordinates and its standard deviation equals to half the extension of the spatial range, which is 10 pixels. Also the pixel contribution is distributed into adjacent spatial cells and orientations bins using trilinear interpolation. This is important to avoid all boundary effects in which the descriptor abruptly changes as a sample shifts smoothly from being within one cell to another or from one orientation to another. Gaussian windowing and trilinear interpolation also increases the robustness of the descriptor against small displacements of the keypoint location.

Finally, HOG descriptors are normalized to increase invariance to illumination changes. As in the Lowe's algorithm, first the 128 elements vector are normalized to unit length. This normalization cancels changes in image contrast. Notice that we do not care about changes in brightness, a constant added to pixel values, because they are suppressed by image gradients. Finally, the descriptor is saturated so that no values over 0.2 are allowed and again re-normalized to unit length. This final step is done to reduce non-linear illumination changes.

### 3 Elastic bunch graph matching

The main idea of EBGGM is that a novel face pattern can be recognized by first localizing a set of facial landmarks and then measuring the similarity between these landmarks and those extracted from a set of faces of each individual.

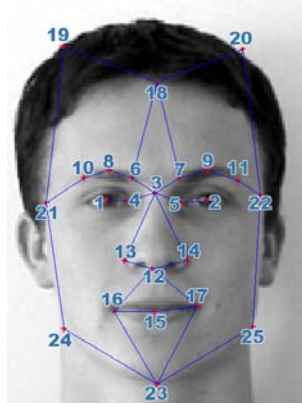


Fig. 2. Face graph and its 25 facial landmarks on a CVL image. The numbers indicate also the search order.

Traditionally, EBGM algorithms use Gabor jets as features for both localization and matching of facial landmarks. In our approach, we replace Gabor coefficients by the HOG descriptors presented in section 2.

Our implementation of EBGM is based on the algorithm developed by Wiskott et al. (6) which was included by the Colorado State University (CSU) as a baseline algorithm for comparison of face recognition algorithms (14). Basically, this HOG-EBGM algorithm can be decomposed into three steps:

- (1) Image normalization
- (2) Creation of face graphs
- (3) Graph matching

The objective of the image normalization step is to reduce variability produced by changes in illumination, scale and rotation.

The next step creates a face graph after the detection of facial landmarks. Of course, the success of the recognition algorithm depends on a good selection of facial landmarks. More precisely, facial landmarks need to be very distinctive between different people and also be relatively easy to detect in a fully automatic system. Our face graph follows the structure proposed in the CSU project (15) with 25 facial landmarks which are shown in Fig. 2, the numbers indicate the search order.

It is important to mention, that not all facial areas contribute equally to face recognition. Several studies (1) have shown that the area around eyes and nose are very important for recognition, for this reason more facial landmarks are placed in these two areas.

Next subsections explain in detail the steps of our HOG-EBGM approach.

### 140 3.1 Image normalization

141 As introduced above, in this paper we assume that the exact location of both  
 142 eyes is known a priori. This simplifies the normalization task since it only  
 143 deals with scale and rotation variations. The normalized face is a  $120 \times 160$   
 144 pixels image in which left and right eyes are located at  $(40, 80)$  and  $(80, 80)$   
 145 pixel coordinates respectively. Normalization in scale is particularly important  
 146 prior to the computation of image gradients.

### 147 3.2 Modeling and location of facial landmarks

148 In EBGM, each face is described by a face graph (FG), composed of strategi-  
 149 cally located keypoints (facial landmarks) and their corresponding descriptors.  
 150 More formally, given a face image, its face graph (FG) is the set of its facial  
 151 landmark coordinates and their associated HOG descriptors:

$$152 \quad FG = \{X_i, J_i = \text{HOG}(X_i), 1 \leq i \leq 25\}$$

153 To automatically locate facial landmarks in new faces, a model for each land-  
 154 mark is needed. These models should account for changes of expression, hair  
 155 styles, illumination, etc. Similar to the original EBGM scheme, we model each  
 156 facial landmark using a set of HOG descriptors called a *bunch*. These HOG  
 157 descriptors are manually obtained from a set of  $N_f$  normalized training faces  
 158 at each particular keypoint. In the sequel,  $\text{fbg}_i(k)$ ,  $1 < i < 25$  and  $1 < k < N_f$ ,  
 159 stands for the HOG descriptor of the  $i$ -th keypoint from the  $k$ -th training face  
 160 (remember that face graphs are composed of 25 keypoints). Finally, we call  
 161 face bunch graph (FBG) to the set of all keypoints models:

$$162 \quad FBG = \{\text{fbg}_i(k), 1 \leq i \leq 25, 1 \leq k \leq N_f\}$$

163 The FBG is used to automatically build new FGs in an iterative process that  
 164 uses the already detected landmarks to reduce the search area. The order in  
 165 which new facial landmarks are searched is set empirically to produce the best  
 166 results. The idea is that since we start from eye locations, the points closer to  
 167 the eyes are detected first.

168 The process to detect the  $i$ -th ( $i > 2$ ) facial landmark is the following:

- 169 (1) Initial estimation of the facial landmark location,  $X_i^s$ . This estimate is  
 170 based on the mean of displacements between the  $i$ -th keypoint and the  
 171  $j$ -th ( $j < i$ ) keypoints. More in detail:  
 172 (a) Let  $d(i, j)$  be the mean displacement between keypoints  $i$  and  $j$  esti-

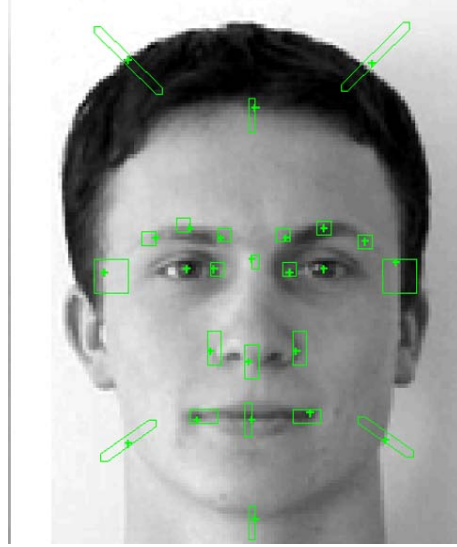


Fig. 3. Search area for each face landmark and detected keypoints.

- 173 mated using the data in the FBG.
- 174 (b) Let  $X_j$  ( $j < i$ ) be the coordinates of the  $j$ -th keypoint which has  
175 been already located.
- 176 (c) For each  $X_j$ , we define  $X_i(j) = X_j + d(i, j)$  as the initial prediction  
177 of the  $i$ -th keypoint based on the  $j$ -th keypoint
- 178 (d) The initial estimate of  $X_i^s = \frac{1}{i-1} \sum_{j<i} X_i(j)$ , i.e. the mean of the  
179 estimates of previous keypoints.
- 180 (2) Calculate the HOG descriptor on the previous location,  $\text{HOG}(X_i^s)$ .
- 181 (3) Compare  $\text{HOG}(X_i^s)$  with the  $\text{fbg}_i(k)$ ,  $0 \leq k \leq N_f$  in the FBG and let:

182 
$$k_{min} = \min_k \|(\text{HOG}(X_i^s) - \text{fbg}_i(k))\|$$

- 183 (4) Define a search area  $S_i$  around  $X_i^s$ . The extent of the search area depends  
184 on the particular keypoint as shown in Fig. 3. We empirically set the  
185 search areas considering the dispersion of the location of facial landmarks  
186 in the FBG for each keypoint.
- 187 (5) Refine the initial estimate of the  $i$ -th keypoint using the descriptor  $\text{fbg}_i(k_{min})$ :

188 
$$X_i = \min_{X \in S_i} \|\text{HOG}(X) - \text{fbg}_i(k_{min})\|$$

189 Note that in this paper, comparison between HOG descriptors is always done  
190 using euclidean distance.



### 191 3.3 Creation and distance between face graphs

192 To compare two face graphs  $FG^k$  and  $FG^l$  from two different faces, we just  
 193 sum up the distances between corresponding keypoint descriptors  $J_i^k$  and  $J_i^l$ :

$$194 \quad D_{kl} = |FG^k, FG^l| = \sum_{i=1}^{25} \|J_i^k - J_i^l\|$$

## 195 4 Experimental setup

196 All the experiments carried out in this paper have been done using three  
 197 sets of images namely: the Yale database (16), a subset of the CVL face  
 198 database (17) and the FERET face database (18). The Yale database con-  
 199 tains 165 frontal images of 15 subjects (11 images/subject) showing different  
 200 expressions and illumination conditions as illustrated in Fig. 4.a. The CVL  
 201 database contains seven images from 114 individuals under uniform illumina-  
 202 tion and different orientations around the vertical axis. In this work, we only  
 203 use the frontal faces of the CVL (3 images/subject). Fig. 4.b shows the selected  
 204 images of the faces from two representative subjects of the CVL database. The  
 205 FERET database contains 3365 full frontal facial images of nearly 1000 sub-  
 206 jects. FERET database images are organized into a gallery set ( $fa$ ) and four  
 207 probe sets ( $fb$ ,  $fc$ ,  $dup1$ ,  $dup2$ ) as illustrated in Fig. 4.c. Using the FERET  
 208 terminology (18) the gallery is the set of known facial images and the probe  
 209 is the set of faces to be identified. The images in sets  $fa$  and  $fb$  were taken in  
 210 the same session with the same camera and illumination conditions but with  
 211 different facial expressions. The  $fc$  images were also taken in the same session  
 212 but using a different camera and different lighting. Finally sets  $dup1$ ,  $dup2$   
 213 are by far the most challenging sets. These images were taken on a later date,  
 214 sometimes years apart, and the photographers sometimes asked the subjects  
 215 to put on their glasses and/or pull their hair back. The reader can see (18) for  
 216 further details.

217 In this paper, we have used the Yale and CVL databases for tuning the param-  
 218 eters of our HOG-EBGM algorithm whereas the FERET database is used as  
 219 a common framework to compare and evaluate our approach respect to other  
 220 face recognition approaches.

221 Finally, it should be said that as a preparation task, we manually marked the  
 222 25 facial landmarks in all images of the Yale and CVL databases. This data  
 223 is used to build the FBG as described in Section 3.2 and in the experiment of  
 224 Section 5.1.1.

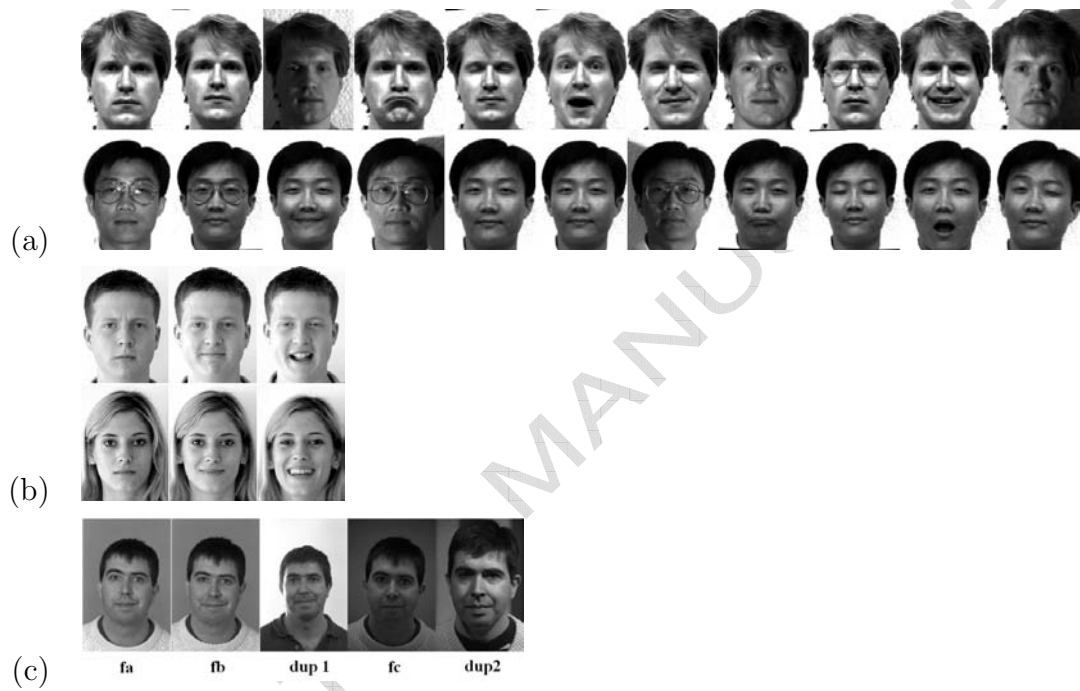


Fig. 4. Examples of subjects from a) Yale database b) CVL database c) FERET database

## 225 5 Experiments and results

226 Two kinds of experiments have been carried out:

- 227 (1) Study the influence and tuning of the HOG-EBGM parameters.
- 228 (2) Comparison of HOG-EBGM against the algorithms provided by the CSU
- 229 Face Identification Evaluation System (14).

### 230 5.1 Tuning of HOG-EBGM parameters

231 In this section we study the influence on the recognition rate of the window  
 232 size of HOG descriptors and the influence of the number of training faces  $N_f$   
 233 of the FBG. The study has been performed using a leave-one-out methodology  
 234 on the CVL and Yale databases. In each experiment round we take one face  
 235 image which is used as a probe set and the remaining images of the database  
 236 are used as the gallery set. Then, the image is recognized using a nearest  
 237 neighbor classifier. This process is repeated for all faces in the database and  
 238 the average is the recognition rate.

#### 239 5.1.1 Influence of the window size of HOG descriptors

240 In this experiment we study the influence of the size of the HOG descriptor  
 241 window. To that end we obtain recognition rates using the Yale and CVL  
 242 databases. To make this study independent of the building process of face  
 243 graphs, all facial landmarks in this experiments are manually marked. The  
 244 window sizes range from  $12 \times 12$  to  $28 \times 28$ . This range takes into consideration  
 245 that the distance between eyes on the normalized-face is always 40 pixels.  
 246 As the reader will notice, window sizes are multiple of 4 because our HOG  
 247 descriptors use a  $4 \times 4$  grid (Fig. 1).

	Size of the HOG Window				
	12 x 12	16 x 16	20 x 20	24 x 24	28 x 28
Yale database	98.2%	97.6%	97.0%	95.8%	94.6%
CVL database	96.2%	98.5%	99.1%	98.8%	98.0%

Table 1

Influence of the feature window size on the recognition rate.

248 The recognition rates are presented in Table 1. It can be seen that  $20 \times 20$  gives  
 249 a maximum when the CVL data is used while the performance for the Yale  
 250 database tends to increase as the window size decreases. This can be explained  
 251 because smaller window sizes make the features more robust to illumination

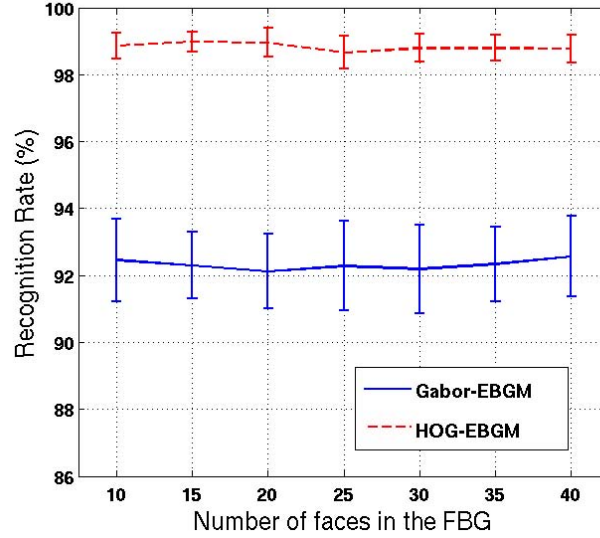


Fig. 5. Performance of Gabor-EBGM and HOG-EBGM for different number of images in the FBG on the CVL database.

changes (which are more noticeable in the Yale database). For this reason, we have selected a  $20 \times 20$  window as a trade off value in the sequel experiments.

#### 5.1.2 Study of the influence of the number of images on the FBG

As described in Section 3.2, our HOG-EBGM approach uses a FBG to automatically detect facial landmarks. The FBG is composed of a set of  $N_f$  training faces which are used to model the facial landmarks. In this experiment we study the influence of the parameter  $N_f$  on the recognition rate. This recognition rate is also compared with the one obtained using standard Gabor-EBGM. Here, the FBG is built using a random set of  $N_f$  training images from the Yale database and the recognition rate is evaluated on the CVL database.

Figure 5 shows the recognition rate when  $N_f$  changes from 10 to 40. Also to check if the algorithm is dependent on the particular set of  $N_f$  images, we make 20 trials with different sets of random images for each value of  $N_f$ . The curves represent the mean recognition rate along with their corresponding standard deviation.

From the results, we can notice that the recognition rate is quite independent of  $N_f$ . This fact is also illustrated in Figure 6. In this figure we show an example of the dispersion of facial landmarks after the 20 trials. We can see that this dispersion is only slightly decreased when the number of faces is set to 40. It can also be seen that the dispersion for the HOG-EBGM is lower than for the Gabor-EBGM which can explain part of the better recognition

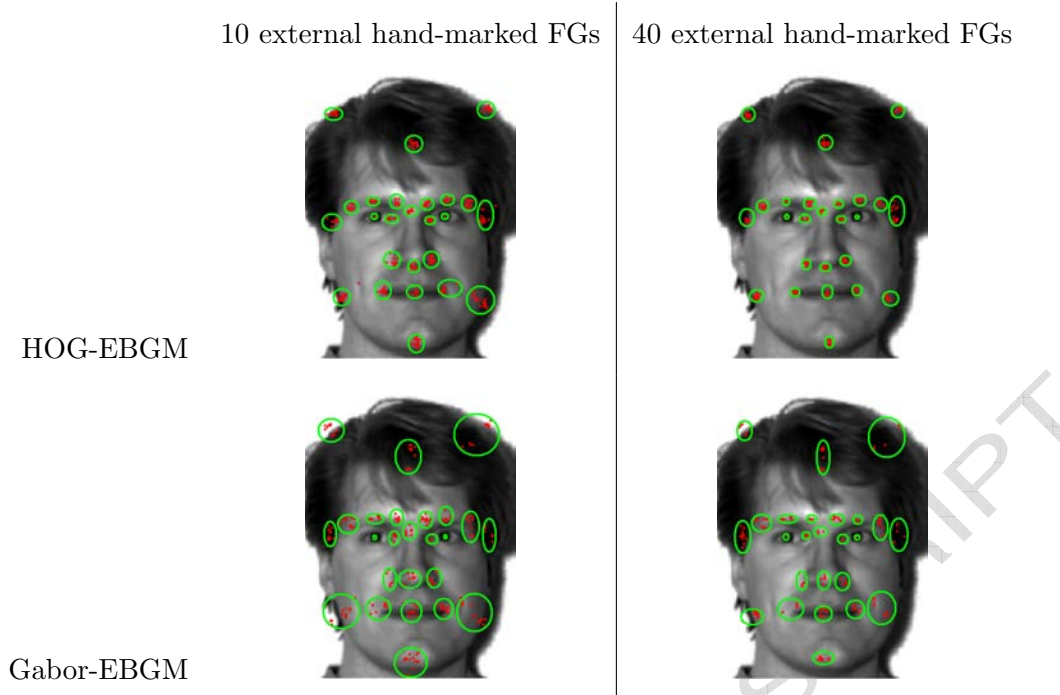


Fig. 6. Dispersion of automatically detected facial landmarks for 10 and 40 images in the FBG. Dispersion is obtained after 20 trials using random sets of images in the FBG.

rate of our approach.

Also from the example in Figure 7 we can see that the FGs located with HOG-EBGM adjust better with their corresponding ideal graphs than the FGs located with Gabor-EBGM. This partly explains the better results in Section 5.2.

From the previous results we set to 10 the number of model images in the FBG for subsequent experiments.

## 5.2 Comparison to other algorithms

In order to evaluate the performance rates of our approach, we have used the the CSU Face Identification Evaluation System (14), which also includes other face recognition algorithms for comparison. The evaluation has been performed on the FERET database using the *fa* set as a gallery and the *fb*, *fc*, *dup1* and *dup2* as probe sets. The algorithms compared in this experiment are:

- Principal Component Analysis (PCA) (2), considering Euclidean and Mahalanobis Cosine distances
- Linear Discriminant Analysis (LDA) (3).

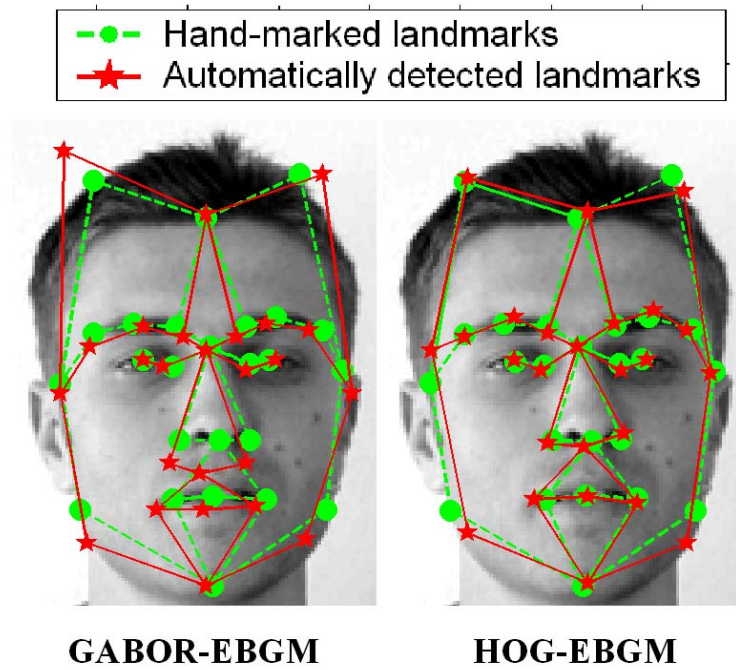


Fig. 7. Comparison of hand-marked and automatic face graphs for both EBGM algorithms.

	PCA Euclidean	PCA Mah. Cos.	LDA LDA	Bayesian MAP	Bayesian ML	Gabor EBGM	HOG EBGM
<i>fafb</i>	74.3%	85.3%	72.1%	81.7%	81.7%	87.3%	<b>95.5%</b>
<i>fafc</i>	5.6%	65.5%	41.8%	35.0%	34.5%	38.7%	<b>81.9%</b>
<i>dup1</i>	33.8%	44.3%	41.3%	50.8%	51.5%	42.8%	<b>60.1%</b>
<i>dup2</i>	14.1%	21.8%	15.4%	29.9%	31.2%	22.7%	<b>55.6%</b>

Table 2

Recognition rate for PCA Euclidean, PCA Mahalanobis Cosine, LDA, Bayesian MAP, Bayesian ML, Gabor-EBGM and HOG-EBGM using the FERET database.

- Bayesian algorithm with variants MAP and ML (19).
- Original Gabor-EBGM algorithm (6).
- Our HOG-EBGM algorithm.

The reader can see the references for implementation details of these algorithms.

We can see in Table 2 the recognition rates obtained for all the experiments. FERET also uses rank curves to compare face recognition algorithms. A rank curve shows for each rank  $k$  the probability that the test face is between the first  $k$  nearest faces. Figure 8 shows the rank curves for the *fafb*, *fafc*, *dup1* and *dup2* experiments.

The results presented in Table 2 and Fig. 8 show that our HOG-EBGM performs better than the other recognition algorithms with all probe sets. This is particularly true for the most difficult probe sets *dup1* and *dup2*.

## 6 Conclusions and future research

This paper presents a new face recognition algorithm based on the well-known EBGM method which replaces Gabor features by HOG descriptors. The recognition results show a better performance of our approach compared to other face recognition approaches using public available databases. This better performance is explained by the properties of HOG descriptors which are more robust to changes in illumination, rotation and small displacements, and to the higher accuracy of the face graphs obtained compared to classical Gabor-EBGM ones.

Future research is focused on improving the classification algorithm (in this work we use nearest neighbor). Also the influence of each particular landmark on the recognition rate will be studied so that landmarks can be properly



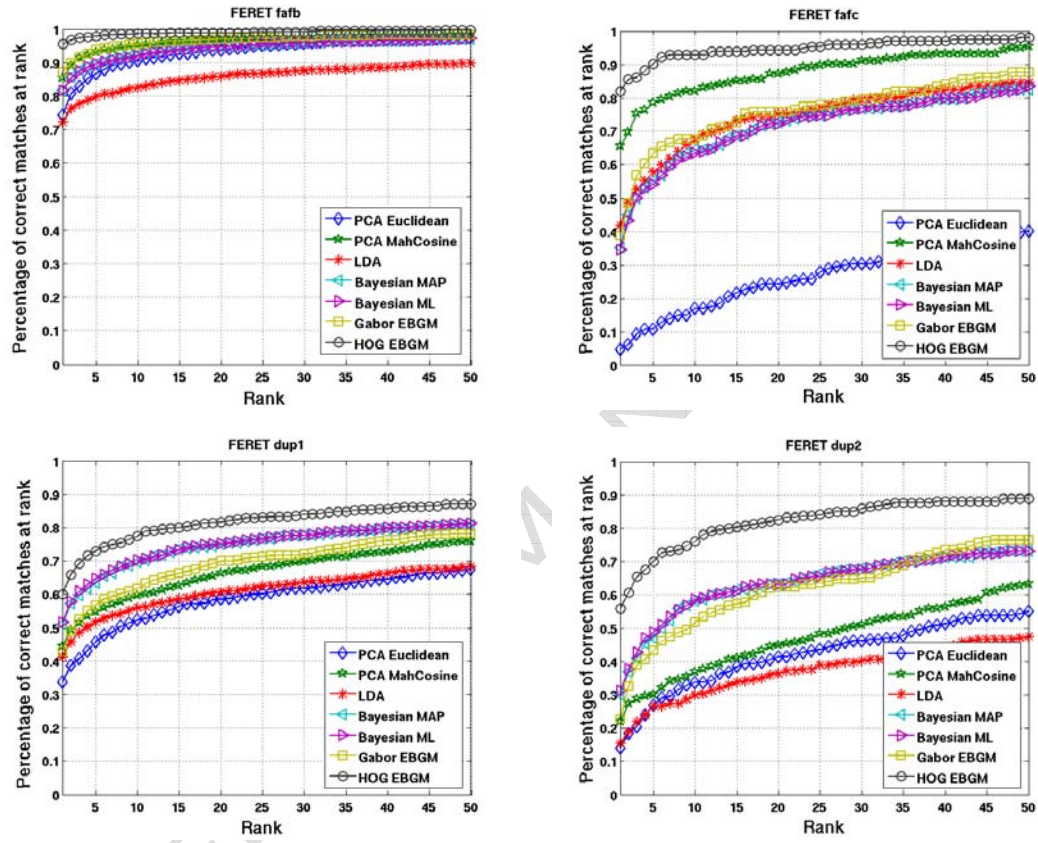


Fig. 8. Rank curves on the FERET database for *fafb*, *fafc*, *dup1* and *dup2* experiments.



316 weighted or resized.

## 317 References

- 318 [1] W. Zhao, R. Chellappa, A. Rosenfeld, P. Phillips, Face recognition: A  
319 literature survey, *ACM Computing Surveys* 35 (4) (2003) 399–458.
- 320 [2] M. Turk, A. Pentland, Eigenfaces for recognition, *Journal of Cognitive*  
321 *Neuroscience* 3 (1) (1991) 71–86.
- 322 [3] P. Belhumeur, J. Hespanh, D. Kriegman, Eigenfaces vs. fisherfaces:  
323 Recognition using class specific linear projection, in: *Proceedings of the*  
324 *4th European Conference on Computer Vision*, Cambridge, UK, 1996,  
325 pp. 45–58.
- 326 [4] J. R. del Solar, P. Navarrete, Eigenspace-based face recognition: a com-  
327 parative study of different approaches, *IEEE Transactions on Systems,*  
328 *Man and Cybernetics* 35 (3) (2005) 315–325.
- 329 [5] T. Cotes, G. J. Edwards, C. J. Taylor, Active appearance models, *IEEE*  
330 *Transactions on Pattern Analysis and Machine Intelligence* 23 (6) (2001)  
331 681–685.
- 332 [6] L. Wiskott, J. M. Fellous, N. Kruger, C. Malsburg, Face recognition  
333 by elastic bunch graph matching, Tech. rep., Ruhr-Universitat Bochum  
334 (April 1996).
- 335 [7] H. Shin, S. D. Kim, H. C. Choi, Generalized elastic graph matching for  
336 face recognition, *Pattern recognition letters* 28 (9) (2007) 1077–1082.
- 337 [8] M. Bicego, A. Lagorio, E. Grosso, M. Tistarelli, On the use of SIFT fea-  
338 tures for face authentication, in: *Proceedings of the International Confer-*  
339 *ence on Computer Vision and Pattern Recognition Workshop*, New York,  
340 2006.
- 341 [9] D. Lowe, Distinctive image features from scale-invariant keypoints, *Inter-*  
342 *national journal of Computer Vision* 60 (2) (2004) 91–110.
- 343 [10] H. Bay, T. Tuytelaars, L. V. Gool, SURF: Speeded up robust features, in:  
344 *Proceedings of the 9th European Conference on Computer Vision*, Graz,  
345 Austria, 2006.
- 346 [11] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors,  
347 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (10)  
348 (2005) 1615–1630.
- 349 [12] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection,  
350 in: *Proceedings of the 9th European Conference on Computer Vision*, San  
351 Diego, CA, 2005.
- 352 [13] P. Viola, M. Jones, Rapid object detection using a boosted cascade of  
353 simple features, in: *Proceedings of the International Conference on Com-*  
354 *puter Vision and Pattern Recognition*, Hawaii, 2001.
- 355 [14] D. S. Bolme, J. R. Beveridge, M. Teixeira, B. Draper, The CSU face  
356 identification evaluation system: Its purpose, features, and structure, in:

- 357        Proceedings of the International Conference on Computer Vision Sys-  
358        tems, Graz, Austria, 2003, pp. 304–313.
- 359 [15] D. Bolme, Elastic bunch graph matching, Ph.D. thesis, Colorado State  
360        University, Fort Collins, Colorado (2003).
- 361 [16] Yale database, <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.
- 362 [17] P. Peer, CVL Face database, University of Ljubjana, <http://www.fri.uni-lj.si/en>.
- 363
- 364 [18] J. P. Phillips, H. Moon, S. Rizv, P. J. Rauss, The FERET evaluation  
365        methodology for face-recognition algorithms, *IEEE Transactions on Pat-*  
366        *tern Analysis and Machine Intelligence* 22 (10) (2000) 1090–1104.
- 367 [19] B. Moghaddam, C. Nastar, A. Pentland, A bayesian similarity measure  
368        for direct image matching, in: *Proceedings of the International Conference*  
369        *on Pattern Recognition*, Vol. 2, Vienna, Austria, 1996, pp. 350–358.