Face Recognition with Patterns of Oriented Edge Magnitudes

Ngoc-Son Vu[1,2] and Alice Caplier[2]

[1] Vesalis Sarl, Clermont Ferrand, France
[2] Gipsa-lab, Grenoble INP, France

Abstract. This paper addresses the question of computationally inexpensive yet discriminative and robust feature sets for real-world face recognition. The proposed descriptor named Patterns of Oriented Edge Magnitudes (POEM) has desirable properties: POEM (1) is an oriented, spatial multi-resolution descriptor capturing rich information about the original image; (2) is a multi-scale self-similarity based structure that results in robustness to exterior variations; and (3) is of low complexity and is therefore practical for real-time applications. Briefly speaking, for every pixel, the POEM feature is built by applying a self-similarity based structure on oriented magnitudes, calculated by accumulating a local histogram of gradient orientations over all pixels of image cells, centered on the considered pixel. The robustness and discriminative power of the POEM descriptor is evaluated for face recognition on both constrained (FERET) and unconstrained (LFW) datasets. Experimental results show that our algorithm achieves better performance than the state-of-the-art representations. More impressively, the computational cost of extracting the POEM descriptor is so low that it runs around 20 times faster than just the first step of the methods based upon Gabor filters. Moreover, its data storage requirements are 13 and 27 times smaller than those of the LGBP (Local Gabor Binary Patterns) and HGPP (Histogram of Gabor Phase Patterns) descriptors respectively.

1 Introduction

Good pattern representation is one of key issues for all pattern recognition systems. In face recognition, a good representation is one which minimizes intra-person dissimilarities whilst enlarging the margin between different people. This is a critical issue, as variations of pose, illumination, age and expression can be larger than variations of identity in the original face images. For real-world face recognition systems we also believe that a good representation should be both fast and compact: if one is testing a probe face against a large database of desirable (or undesirable) target faces, the extraction and storage of the face representation has to be fast enough for any results to be delivered to the end user in good time. In this paper, we propose a novel feature descriptor named Patterns of Oriented Edge Magnitudes (POEM) for robust face recognition, a descriptor which we argue satisfies these criteria. Experimental results on both

K. Daniilidis, P. Maragos, N. Paragios (Eds.): ECCV 2010, Part I, LNCS 6311, pp. 313–326, 2010.

FERET and LFW databases show that POEM method achieves comparable and better performance when compared with state-of-the-art representations. More impressively, the runtime required to extract our descriptor is around 20 times faster than that of even the first step of methods based upon Gabor filters.

We briefly discuss related work in Section 2, describe our method in Section 3. Section 4 details the use of POEM for face recognition. Experimental results are presented in Section 5 and conclusions are given in Section 6.

2 Related Work

There is an extensive literature on local descriptors and face recognition. We refer readers to [1, 2] for an in-depth survey, and describe here those high-performing algorithms which are most relevant to our work [3–8].

Local descriptors [1, 7, 9] are commonly employed for many real-world applications because they can be computed efficiently, are resistant to partial occlusion, and are relatively insensitive to changes in viewpoint. Mikolajczyk and Schmid [1] recently evaluated a variety of local descriptors and identified the SIFT (Scale-invariant feature transform) [9] algorithm as being the most resistant to common image deformations. As a dense version of the dominating SIFT feature, HOG [6] has shown great success in object detection and recognition [6, 10] although has not seen much use in face recognition.

For the specific problem of face recognition, there are also many representational approaches, including subspace based holistic features and local appearance features. Heisele *et al.* [11] compared local and global approaches and observed that local systems outperformed global systems for recognition rates larger than 60%. Due to increasing interest, in recent surveys stand-alone sections have been specifically devoted to local methods.

Two of the most successful local face representations are Gabor features [3, 4, 8, 12–15] and Local Binary Patterns (LBP) [5, 13, 16, 17]. Gabor filters, which are spatially localized and selective to spatial orientations and scales, are comparable to the receptive fields of simple cells in the mammalian visual cortex [8]. Due to their robustness to local distortions, Gabor features have been successfully applied to face recognition. Both the FERET evaluation and the FVC2004 contests have seen top performance from Gabor feature based methods. Recently, Pinto *et al.* [14, 15] use V1-like and V1-like+, the Gabor filter based features, as face representation and report good recognition performance on the unconstrained LFW set. Gabor features are typically calculated by convolving images with a family of Gabor kernels at different scales and orientations, which is a costly stage. Despite recent attempts at speeding this process up (e.g. the Simplified Gabor Wavelets of Choi *et al.* [18]) the process of extracting these features is still prohibitive for real-time applications.

More recently, the spatial histogram model of LBP has been proposed to represent visual objects, and successfully applied to texture analysis [19], and face recognition [5]. LBP is basically a fine-scale descriptor that captures small texture details, in contrast to Gabor features which encode facial shape and appearance over a range of coarser scales. By using LBP, Ahonen *et al.* [5] have reported impressive results on the FERET database. Several variants of LBP are also presented and successfully applied to different applications, such as Center-Symmetric LBP (CSLBP) [20], Three-Patch LBP (TPLBP), Four-Patch LBP (FPLBP) [16], etc.

A combination approach was introduced by Zhang *et al.* [3] extending LBP to LGBP by introducing multi-orientation and multi-scale Gabor filtering before using LBP and impressively improved the performance when compared with pure LBP. In a similar vein, they further proposed HGPP [4] combining the spatial histogram and the Gabor phase information encoding scheme. In [13], a model fusing the multiple descriptor sets is presented with very high performance on constrained datasets. More recently, Wolf *et al.* [17] combine LBP, TPLBP, FPLBP, Gabor and SIFT with different similarity measures, showing promising results on the LFW set.

These combination methods try to bring the advantages of LBP and Gabor filters together, but they also bring the disadvantages of Gabor based systems; namely computational cost and storage requirements.

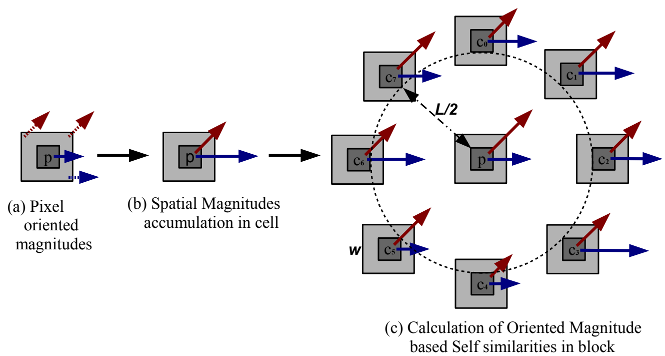

Fig. 1. Main steps of POEM feature extraction

The aim of this study is to find a feature descriptor that can inherit various good properties from existing features but with low computational cost. We propose applying the LBP based structure on oriented magnitudes to build a novel descriptor: Patterns of Oriented Edge Magnitudes (POEM). Briefly speaking, in order to calculate the POEM for one pixel, the intensity values in the calculation of the traditional LBP are replaced by the gradient magnitudes, calculated by accumulating a local histogram of gradient directions over all pixels of a spatial patch ("*cell*"). Additionally, these calculations are done across different orientations. We use the terms *cell* and *block*, as in [6], but with a slightly different meaning. Cells (big squares in Figure 1a) refer to spatial regions around the current pixel where a histogram of orientation is accumulated and assigned to the cell central pixel. Blocks (circular in Figure 1c) refer to more extended spatial regions, on which the LBP operator is applied. Note that our use of oriented magnitudes is also different from that in [6] where HOG is computed in dense grids and then is used as the representation of cell. On the contrary, in POEM, for *each pixel*, a local histogram of gradient over all pixels of cell, centered on the considered pixel, is used as *the representation of that pixel.* Similarly, the term *pattern* in POEM is not as *local* as in the conventional LBP based methods. LBP methods often calculate the self-similarity within a small neighborhood while the block used in POEM is rather extended (see details in sections 3.2 and 5.1).

In combination approach [3], Gabor filters are first used for capturing large scale information and LBP operator is then applied for encoding the small details. On the contrary, POEM first characterizes object details in small scale and then uses the LBP based structure to encode information over larger region.

3 POEM Descriptor

Similar features have seen increasing use over the past decade [6, 7, 9]; the fundamental idea being to characterize the local object appearance and shape by the distribution of local intensity gradients or edge directions. We further apply the idea of self-similarity calculation from LBP-based structure on these distributions since we find that combining both the edge/local shape information and the relation between the information in neighboring cells can better characterize object appearance. As can be seen in Figure 1, once the gradient image is computed, the next two steps are assigning the cell's accumulated magnitudes to its central pixel, and then calculating the block self-similarities based on the accumulated gradient magnitudes by applying the LBP operator.

3.1 POEM Feature Extraction in Detail

The first step in extracting the POEM feature is the computation of the gradient image. The gradient orientation of each pixel is then evenly discretized over 0-π (*unsigned* representation) or 0-2π (*signed* representation). Thus, at each pixel, the gradient is a 2D vector with its original magnitude and its discretized direction (the blue continuous arrow emitting from pixel $\mathbf{p}$ in Figure 1a).

The second step is to incorporate gradient information from neighbouring pixels (the discontinuous arrows in Figure 1a) by computing a local histogram of gradient orientations over all cell pixels. Vote weights can either be the gradient magnitude itself, or some function of the magnitude: we use the gradient magnitude at the pixel, as in [6]. At each pixel, the feature is now a vector of m values where m is the number of discretized orientations (number of bins).

Finally, we encode the accumulated magnitudes using the LBP operator within a block. The original LBP operator labels the pixels of an image by thresholding the 3×3 neighborhood surrounding the pixel with the intensity value of central pixel, and considering the sequence of 8 result bits as a number (as shown in Figure 2). Only uniform patterns, which are those binary patterns that have at most 2 transitions from 0 to 1, are typically used to accelerate the method.

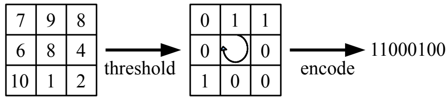

Fig. 2. LBP operator

We apply this procedure on the accumulated gradient magnitudes and across different directions to build the POEM. Firstly, at the pixel position p, a POEM feature is calculated for each discretized direction θ_i:

$$POEM^{\theta_i}_{L,w,n}(p) = \sum_{j=1}^{n} f(S(I^{\theta_i}_p, I^{\theta_i}_{c_j}))2^j, \tag{1}$$

where I_p, I_{c_j} are the accumulated gradient magnitudes of central and surrounding pixels p, c_j respectively; $S(.,.)$ is the similarity function (e.g. the difference of two gradient magnitudes); L, w refer to the size of blocks and cells, respectively; n, set to 8 by default in this paper, is number of pixels surrounding the considered pixel p; and f is defined as:

$$f(x) = \begin{cases} 1 \; if \; x \geq \tau, \\ 0 \; if \; x < \tau, \end{cases} \tag{2}$$

where the value τ is slightly larger than zero to provide some stability in uniform regions, similar to [16].

The final POEM feature set at each pixel is the concatenation of these unidirectional POEMs at each of our m orientations:

$$POEM_{L,w,n}(p) = \{POEM^{\theta_1}, ..., POEM^{\theta_m}\}, \tag{3}$$

3.2 Properties of POEM

We discuss here the good properties of this novel descriptor for object representation postponing the question of complexity until Section 5.4. For each pixel, POEM characterizes not only local object appearance and shape, but also the relationships between this information in neighboring regions. It has the following properties:

- POEM is an oriented feature. Since the number of discretized directions can be varied, POEM has the ability to capture image information in any direction and is adaptable for object representation with different levels of orientation accuracy.
- Computed at different scales of cells and blocks, POEM is also a spatial multi-resolution feature. This enables it to capture both local information and more global structure.
- Using gradient magnitudes instead of the pixel intensity values for the construction makes POEM robust to lighting variance. In [21, 22], edge magnitudes have been shown to be largely insensitive to lighting.

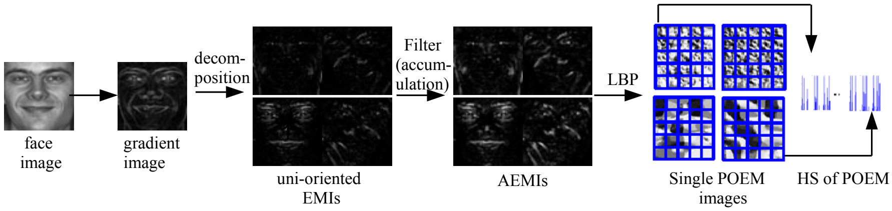

Fig. 3. Implementation of POEM for face description

– The oriented magnitude based representation contains itself the relation between cell pixels. POEM further calculates dissimilarities between cells and therefore has the ability to capture multi-scale self-similarities between image regions. This makes POEM robust to exterior variations, such as local image transformations due to variations of pose, lighting, expression and occlusion that we frequently find when dealing with faces.

Patch-based or multi-block LBP [16] also considers relationships between regions in a similar way to our POEM descriptor. However the richer information coming from the use of gradients at multiple orientations gives us greater descriptive power, and a greater insensitivity to lighting variations.

4 Face Recognition Based on POEM

For face recognition, we use a similar procedure to that described in [5], except that each pixel is characterized with the POEM features instead of a LBP code (cf., Figure 3).

POEM Histogram Sequences for Face Recognition

In practice, the Oriented Edge Magnitude Image (oriented EMI) is first calculated from the original input image (section 3.1) and divided into m uni-oriented EMIs through gradient orientations of pixels. Note that the pixel value in uni-oriented EMIs is gradient magnitude. For every pixel on uni-oriented EMIs, its value is then replaced by the sum of all values in the cell, centered on the current pixel. These calculations are very fast (using the advantage of integral image [23]). Result images are referred to accumulated EMIs (AEMIs). LBP operators are applied on these AEMIs to obtain the POEM images (Figure 3). In order to incorporate more spatial information into the final descriptor, the POEM images are spatially divided into multiple non-overlapping regions, and histograms are extracted from each region. Similar to [5, 16], only *uniform* POEM codes are used. Finally, all the histograms estimated from all regions of all POEMs are concatenated into a single histogram sequence (POEM-HS) to represent the given face.

Given two histogram sequences of POEM representing two face images, we use the chi-square distance between histograms [5] to measure the similarity between two images.

5 Experiments and Discussions

In this section, we conduct comparison experiments on two face databases, FERET (controlled variations) [24] and LFW (unconstrained environments) [25], in order to validate the efficiency of the proposed descriptor for face recognition.

5.1 Parameter Evaluation

In this section we consider how the parameters of POEM influence its final performance. Parameters varied include the number m and type (unsigned or signed) of orientations, the cell size $(w * w)$, and block size $(L * L)$. As for the cell/block geometry, two main geometries exist: rectangular and circular. In this paper we use circular blocks including bilinear interpolation for the values since they provide the relation between equidistant neighboring cells [5]. Square cells are used, meaning that pixel information is calculated using its neighborhood in a square patch.

The experiments checking the effects of parameters are conducted on the FERET face database, following the standard evaluation protocol: Fa containing 1196 frontal images of 1196 subjects is used as Gallery, while Fb (1195 images of expression variations), Fc (194 images of illumination variations), Dup I & Dup II(722 & 234 images taken later in time) are the Probe sets.

The facial images of FERET are first cropped and aligned using the given coordinates of two eyes. We roughly fixe the width and height of face about two times of distance between the centers of the two eyes, and then resize image to 110x110 pixels (cf., the first image in Figure 3). We do not use any other particular face mask, such as in [5] for example. In our experiments, we divide each image into 10x10 non overlapping patches. Since this paper concentrates on the feature sets, we use a simple nearest neighbor classifier to calculate the recognition rates and consider classifier choice beyond the scope of the current paper. But we believe that better classifier could enhance recognition performance.

Experiment 1, concerning the number of orientations and signed/ unsigned representation. Nearly six hundred cases are considered, recognition rates are calculated on 3000+ face images with different parameters: $L = \{5, 6, 7, 8, 9, 10, 11\}$, $w = \{3, 4, 5, 6, 7, 8\}$, the numbers of discretized orientations are $m = \{2, 3, 4, 5, 6, 7\}$ in the case of unsigned representation, and are doubled to $m = \{4, 6, 8, 10, 12, 14\}$ in the case of signed representation. Cells can overlap, notably when blocks are smaller than cells, meaning that each pixel can contribute more than once.

For each Probe set, the average rates are calculated over different numbers and types of orientation. Figure 4 shows the recognition rates obtained on Probe sets Fb, Fc, Dup1, and Dup2. The average recognition rates obtained on Probe set Fb (as shown in the Figure 4 a) are arround 96.5%, representing an improvement of about 3.5% in comparison with LBP [5].

Considering the question of using a signed or an unsigned representation, we find similar results to [6], in that including signed gradients decreases the

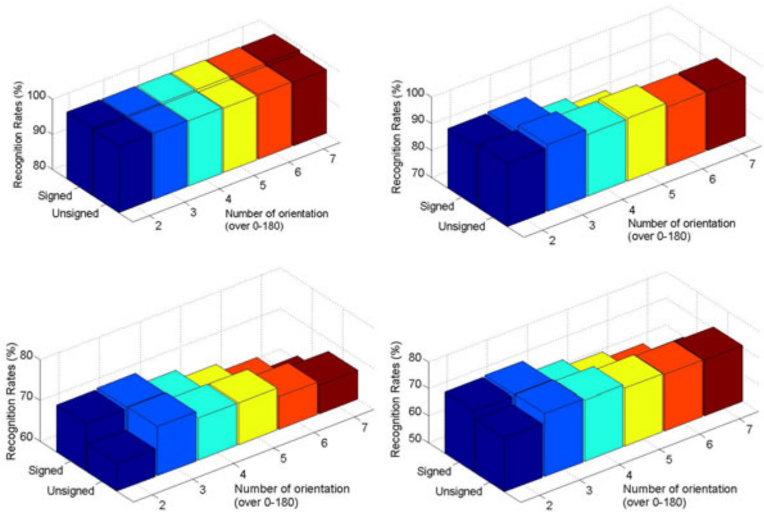

Fig. 4. Recognition rates obtained with different numbers of orientations on Probe sets: Fb (a), Fc (b), Dup1 (c), and Dup2 (d). These rates are calculated by averaging recognition rates with different sizes of cell/block.

performance of POEM even when the data dimension is doubled to preserve more original orientation resolution. For face recognition, POEM provides the best performance with only 3 unsigned bins. This should be noted as one advantage of POEM since the data dimension for face description is not greatly increased as in LGBP or HGPP [3, 4]. It is clear from Figure 4 (c,d) that using too many orientations degrades significantly the recognition rates on Dup1 and Dup2 sets. This can be explained by the fact that increasing the number of orientation bins makes POEM more sensitive to wrinkles appearing in face with time.

Summarizing, the number m and signed/unsigned orientations do not affect the recognition performance in the case of expression variations (Fb set); the unsigned representation is more robust than signed representation, notably to lighting; using few orientations (1, 2) is not enough to represent face information but too many (more than 3) makes POEM sensitive to aging variations. Thus, the best case is 3 unsigned bins.

Experiment 2, concerning the size of cells and blocks. Average recognition rates of all four Probe sets are first calculated with different sizes of cells and blocks with 3 unsigned bins of orientation discretization. As can be seen from Figure 5, using POEM built on 10x10 pixel blocks with histogram of 7x7 pixel cells provides the best performance.

To verify the correctness of these parameters, we further calculate the average rates across cell sizes and across block sizes, meaning that these parameters are now considered independent. Also, in this test, both 10x10 pixel block and 7x7 pixel cell

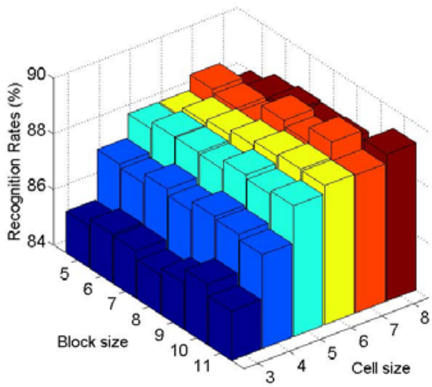

Fig. 5. Recognition rates as the cell and block sizes change

performed the best. This procedure has been repeated with different numbers of orientation bins, and the same optimal parameters have been obtained.

In conclusion, the optimal POEM parameters for face recognition are: unsigned representation with 3 bins, built on 10x10 pixel blocks and 7x7 pixel cells.

5.2 Results Based Upon the FERET Evaluation Protocol

We consider the FERET'97 results [24], results of the LBP [5], HGPP [4], LGBPHS [3], and more recent results in [12, 13, 26]. These results, to the best of our knowledge, are the state-of-the-art with respect to the FERET dataset.

As can be seen from Table 1, in comparison with the conventional LBP and HOG (the perfomance of HOG for face recognition is reported in [26]), our POEM descriptor is much more robust to lighting, expression & aging, illustrated by significant improvements in recognition rates for all probe sets. While compared with LGBP and HGPP, reported as being the best performing descriptors on FERET database, POEM provides comparable performance for the probe sets Fb and Fc. When we consider the more challenging probe sets Dup1 and Dup2, POEM outperforms LGBP and is comparable to HGPP.

Concerning the results of [12, 13], they are only suitable for very limited reference since they are obtained by using a more complex classification phase, and we wish to concentrate upon the performance of the descriptor rather than the classifiers. In [12], Zou *et al.* use Gabor jets as local facial features which are compared using normalized inner products at different scales, and results are combined using the Borda Count method. Moreover, Zou *et al.* do not use only the pure face area, as defined in [27]. In [13], Tan and Triggs fuse two feature sets, Gabor & LBP, and use a complex dimensionality reduction and classification phase, PCA & Kernel DCV. Their method suffers from the disadvantage that adding a new individual to the gallery requires recalculating all existing coefficients: PCA coefficients of Gabor & LBP features, and the KDCV coefficients of the fused features.

Table 1. Recognition rate comparisons with other state-of-the-art results tested with Feret evaluation protocol

Methods	Fb	Fc	Dup1	Dup2
LBP [5]	93.0	51.0	61.0	50.0
LGBPHS [3]	94.0	97.0	68.0	53.0
HGPP [4]	97.6	98.9	77.7	76.1
HOG [26]	90.0	74.0	54.0	46.6
POEM	**97.6**	**96**	**77.8**	**76.5**
Retina filter [28] + POEM	**98.1**	**99**	**79.6**	**79.1**
Results of [12]	*99.5*	*99.5*	*85*	*79.5*
Results of [13]	*98*	*98*	*90*	*85*

We further employ the real-time retina filtering presented by Vu and Caplier in [28] as preprocessing step since this algorithm, as pointed out by authors, not only removes the illumination variations but also enhances the image edges, upon which our POEM is constructed. It is clear from Table 1 that the retina filter enhances the performance of POEM, especially for the probe set Fc.

5.3 Results on LFW Dataset

In order to test the performance of the POEM descriptor across different databases, we duplicate these experiments on another well-known dataset, LFW [25], containing 13233 face images of 5749 individuals. This database is described as "unconstrained", meaning that face images are subject to a large range of "natural" variations. The operational goal of this set differs from above FERET database; it is aimed at studying the problem of face pair matching (given two face images, decide whether they are from the same person or not). We follow the standard procedure described in [25] and report the mean classification accuracy $\pm$ standard error computed from 10 folds of the "Image-Restricted View 2" portion of LFW set.

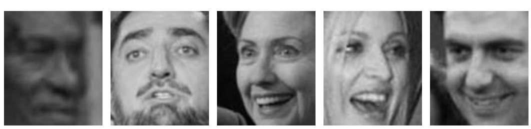

Fig. 6. Examples of LFW images used in our tests

As mentioned above, the goal of the current paper is to demonstrate the efficiency of the novel descriptor, not to compete in the LFW challenge. We therefore report the obtained results using POEM descriptor in a simple threshold-on-descriptor-distance classification context [16], meaning that for each test fold, an optimal threshold giving the highest separation score on the 5400 examples of the training set is chosen and then is used to calculate the classification accuracy for the 600 examples of the test set. We only compare our results with other

Table 2. Recognition results of differents methods on LFW set, Image-Restricted Training, View 2

Reference	Descriptors (similarity measure)	Performance
Pinto2008 [14]	V1-like	0.6421 ± 0.0069
	V1-like+	0.6808 ± 0.0045
Wolf2009 [17]	LBP Euclidean/SQRT	0.6824/0.6790
	Gabor Euclidean/SQRT	0.6849/0.6841
	TPLBP Euclidean/SQRT	0.6926/0.6897
	FPLBP Euclidean/SQRT	0.6818/0.6746
	SIFT Euclidean/SQRT	0.6912/0.6986
	All combined	0.7521 ± 0.0055
This paper	**POEM**	**0.7400 ± 0.0062**
	POEM Flip	**0.7542 ± 0.0071**

descriptor-based results and refer readers to [29] for further algorithm classifiers reported on the LFW dataset.

In our experiments, the LFW gray images aligned automatically by Wolf *et al.* [17] are used and cropped to 100 x 116 pixels around their center. As is clear from the Figure 6, there is significant pose variation within this dataset. In order to address this we flip image 1 of each pair on the vertical axis, and take the smaller of the two histogram distances as our measure. This simple preprocessing step improves recognition rates and is refered to as POEM-Flip in following results. Because of the poor quality of the images in the LFW dataset, retina filtering does not improve the recognition results. With low quality images, the retina filter enhances the image contours and removes illumination variations but also enhances image artifacts (such as those arising from compression). The similar performance is obtained in both cases that the retina filter is used or not. Otherwise, we do not employ any other preprocessing technique.

It is clear from Table 2 that POEM method outperforms all other competing descriptors: LBP, TPLBP, FPLBP, Gabor filters and SIFT. When compared with these descriptors, the POEM based method represents around 20% reduction in classification error. Our POEM-flip mean recognition rate 75.42% is better than that of the "combined" method of [17]. It is worth noting that Wolf *et al.* [17] combine 10 descriptor/mode scores using SVM classification. The results in [14], based upon Gabor filters, are much worse than ours.

5.4 A Consideration of Computational Cost

In this section we compare the complexity of POEM with two of the most widely used descriptors for face recognition: LBP and Gabor wavelet based methods. Considering the pure one-LBP-operator method, POEM based face recognition requires a computational complexity which is 3 times higher (the calculation of integral gradient image is very fast when compared to the calculation of POEM features and the construction of POEM-HS) but at the same time, there are

Table 3. Runtime required to extract the whole POEM descriptor and the initial step of Gabor based feature extraction. Calculated using the implementation in Matlab, these times are only suitable for rough comparisons of computational complexity.

Methods	Times (seconds)
Convolution with 40 Gabor kernels	0.4349
POEM extraction	0.0191

remarkable improvements in recognition rates on the FERET database (+5%, +45%, +16.6% and +26.5% for the probe sets Fb, Fc, Dup1 and Dup2, respectively). And on the LFW set, POEM method also outperforms other variants of LBP, TPLBP and FPLBP.

When we consider Gabor filter based descriptors, only the runtime required for the convolution of the image with the family of Gabor kernels (8 orientations and 5 scales) is necessary. From Table 3, we see that the computation of the whole POEM descriptor is about 23 times faster than that of just this first step of Gabor feature extraction.

We do not calculate here the time required to extract SIFT descriptor and do not compare directly it to POEM, but as argued in [20], SIFT is about 3 times slower than 3×3 grid Center-Symmetric LBP (CSLBP), a variant of LBP (3×3 grid CSLBP means that the descriptor is obtained concatenating the histogram of CSLBP features over grid of 3×3). Thus it seems that POEM and SIFT have the similar time complexity. However, for face recognition, POEM clearly ourperforms SIFT, representing about 20% reduction in classification error on LFW set. Retina filtering is a linear and real-time algorithm. Its calculation time is about 1/5 of that required to extract POEM.

Considering data storage requirements, for a single face, the size of a complete set of POEM descriptors is 13 and 27 times smaller than that of LGBP and HGPP (LGBP calculates LBP on 40 convolved images while HGPP encodes both real and imaginary images). Note that these comparisons are roughly done considering all 256 patterns of our POEM features. However, in this paper, we use only 59 uniform POEMs, meaning that the size of POEM descriptors used here is 58 and 116 times smaller that of LGBP and HGPP (these ones use all 256 feature values). When compared to the "combined" method of Wolf *et al.* [17], the space complexity of POEM descriptor is considerably smaller. For one patch, the size of POEM-HS is 59×3, while the size of method in [17] is $59 \times 2 + 16$ (the size of LBP, TPLBP, and FPLBP per patch are 59, 59 & 16, respectively) + 128 (dimension of SIFT) + size of Gabor based descriptor (which is equal to the size of patch $\times$ number of scales $\times$ number of orientations as in [17, 26]).

Thus we agrue that the POEM descriptor is the first to allow high performance real-time face recognition. Low complexity descriptors provide worse results; whilst representations based upon multiple feature types can achieve similarly high performance but are too slow for real-time systems.

6 Conclusion

By applying the LBP operator on accumulated edge magnitudes across different directions, we have developed a novel descriptor for face representation which has several desirable features. It is robust to lighting, pose and expression variations, and is fast to compute when compared to many of the competing descriptors. We have shown that it is an effective representation for face recognition in both constrained (FERET) and unconstrained (LFW) face recognition tasks outperforming all other purely descriptor based methods. This high performance coupled with the speed of extraction suggests that this descriptor is a good candidate for use in real-world face recognition systems.

Future work will involve testing the POEM descriptor in a broader range of computer vision tasks, such as face detection and object recognition. We will also investigate the use of more powerful classifiers alongside the POEM descriptor within the face recognition domain.

References

1. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE TPAMI 27, 1615–1630 (2005)
2. Zhao, W., Chellappa, R., Corporation, S., Rosenfeld, A., Phillips, P.J.: Face recognition: A literature survey. ACM Computing Surveys (2000)
3. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. In: ICCV, vol. 1, pp. 786–791 (2005)
4. Zhang, B., Shan, S.S., Chen, X.: Gao: Histogram of gabor phase patterns (hgpp): A novel object representation approach for face recognition. IEEE Trans. on Image Processing 16, 57–68 (2007)
5. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
6. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR, Washington, DC, USA, pp. 886–893. IEEE Computer Society, Los Alamitos (2005)
7. Ke, Y., Sukthankar, R.: Pca-sift: A more distinctive representation for local image descriptors. In: CVPR (2004)
8. Liu, C., Wechsler, H.: Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Trans. on Image Processing 11, 467–476 (2002)
9. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. Journal of Computer Vision 60, 91–110 (2004)
10. Felzenszwalb, P., McAllester, D., Ramanan, D.: A discriminatively trained, multiscale, deformable part model (2008)
11. Heisele, B., Ho, P., Wu, J., Poggio, T.: Face recognition: component-based versus global approaches. Computer Vision and Image Understanding 91, 6–21 (2003)
12. Zou, J., Ji, Q., Nagy, G.: A comparative study of local matching approach for face recognition. IEEE Trans. on Image Processing 16(10), 2617–2628 (2007)

13. Tan, X., Triggs, B.: Fusing gabor and lbp feature sets for kernel-based face recognition. In: Zhou, S.K., Zhao, W., Tang, X., Gong, S. (eds.) AMFG 2007. LNCS, vol. 4778, pp. 235–249. Springer, Heidelberg (2007)
14. Pinto, N., DiCarlo, J., Cox, D.: Establishing good benchmarks and baselines for face recognition. In: Faces in Real-Life Images Workshop in ECCV (2008)
15. Pinto, N., DiCarlo, J., Cox, D.: How far can you get a modern face recognition test set using only simple features? In: CVPR (2009)
16. Wolf, L., Hassner, T., Taigman, Y.: Descriptor based methods in the wild. In: Real-Life Images workshop at ECCV (2008)
17. Wolf, L., Hassner, T., Taigman, Y.: Similarity scores based on background samples. In: ACCV (2009)
18. Choi, W., Tse, S., Wong, K., Lam, K.: Simplified gabor wavelets for human face recognition. Pattern Recognition 41, 1186–1199 (2008)
19. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE TPAMI 24, 971–987 (2002)
20. Heikkila, M., Pietikainen, M., Schmid, C.: Description of interest regions with local binary patterns. Pattern Recognition 42, 432–436 (2009)
21. Chen, H.F., Belhumeur, P.N., Jacobs, D.W.: In: search of illumination invariants. In: CVPR (2000)
22. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.: A study of face recognition as people age. In: ICCV (2007)
23. Viola, P., Jones, M.: Robust real-time face detection. Int. Journal of Computer Vision 57, 137–154 (2004)
24. Phillips, J., Moon, H., Rizvi, S.A., et al.: The feret evaluation methodology for face-recognition algorithms. IEEE TPAMI 22, 1090–1104 (2000)
25. Huang, G.B., Manu Ramesh, T.B., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical report, University of Massachusetts, Amherst (2007)
26. Meyers, E., Wolf, L.: Using biologically inspired features for face processing. Int. Journal of Computer Vision 76, 93–104 (2008)
27. Chen, L., Liao, H., Lin, J., Han, C.: Why recognition in a statistics-based face recognition system should be based on the pure face portion: a probabilistic decision-based proof. Pattern Recognition 34(5), 1393–1403 (2001)
28. Vu, N., Caplier, A.: Illumination-robust face recognition using the retina modelling. In: ICIP (2009)
29. LFWresults: `http://vis-www.cs.umass.edu/lfw/results.html`