

RESEARCH

Open Access

Explore semantic pixel sets based local patterns with information entropy for face recognition

Zhenhua Chai^{1*}, Heydi Mendez-Vazquez², Ran He¹, Zhenan Sun¹ and Tieniu Tan¹

Abstract

Several methods have been proposed to describe face images in order to recognize them automatically. Local methods based on spatial histograms of local patterns (or operators) are among the best-performing ones. In this paper, a new method that allows to obtain more robust histograms of local patterns by using a more discriminative spatial division strategy is proposed. Spatial histograms are obtained from regions clustered according to the semantic pixel relations, making better use of the spatial information. Here, a simple rule is used, in which pixels in an image patch are clustered by sorting their intensity values. By exploring the information entropy on image patches, the number of sets on each of them is learned. Besides, Principal Component Analysis with a Whitening process is applied for the final feature vector dimension reduction, making the representation more compact and discriminative. The proposed division strategy is invariant to monotonic grayscale changes, and shows to be particularly useful when there are large expression variations on the faces. The method is evaluated on three widely used face recognition databases: AR, FERET and LFW, with the very popular LBP operator and some of its extensions. Experimental results show that the proposal not only outperforms those methods that use the same local patterns with the traditional division, but also some of the best-performing state-of-the-art methods.

1 Introduction

Face recognition is a popular biometric technique, mainly because it is considered as non-intrusive and it can be applied in a wide range of applications such as access control, video surveillance and human computer interaction [1]. Feature extraction is one of the most important steps in the face recognition process, but to obtain discriminative and robust features for describing face images is still an open problem [2]. Several methods have been proposed toward this aim, that can mainly be divided into two groups: local feature-based methods and global appearance-based methods [1]. In general, local feature-based methods exhibit a better behavior and have some advantages over the global ones [3-5]. Among existing local descriptors, Gabor wavelet-based methods are one of the best performing, mainly due to their spatial locality

and orientation selectivity [6]. However, although different strategies have been proposed, they are still computationally intensive and consume too much time in feature extraction [7], being not suitable for real-time and mobile applications. On the other hand, histograms of local patterns, such as Local Binary Patterns and its different extensions, which are also very popular local descriptors [8], are very simple and fast to compute.

The Local Binary Patterns (LBP) operator was first proposed for texture classification and was then applied to face recognition using a regular regions division [9]. Many extensions of the original operator have appeared afterwards [10-16]; however most of them have focused on obtaining more discriminative descriptors, while few methods have been proposed to get a more robust division strategy.

Recently, the semantic pixel set-based LBP (spsLBP) [17] was proposed for this aim. By clustering the pixels in an image region into a number of sets according to their semantic meanings instead of using a regular division, it

*Correspondence: zhchai@nlpr.ia.ac.cn

¹National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science, P.O. Box 2728, Beijing 100190, People's Republic of China
Full list of author information is available at the end of the article

makes better use of the spatial information when constructing the local histograms. It was shown in [17] that this strategy can alleviate to some extent the pixel-shifting problem caused by some face deformations like variations in expression. However, only the original LBP operator was tested with the proposed strategy in [17], while more robust LBP variants can be used for improving the overall performance.

In this paper, we aim at extending the proposal in [17] to a more general framework, in which more robust local operators can be applied, such as Local Ternary Patterns (LTP) and Three-Patch LBP (TP-LBP). Moreover, we believe that the amount of information in different face regions is different, then using a fixed number of sets for all regions, like in [17], could not be appropriate. Hence, a different number of sets should be used in different regions according to their specific information quantity. Taking this into account, we propose in this paper a method for automatically learning the number of sets in which each region should be divided, by using information entropy. When including more sets, the feature vector dimensionality increases, so a dimensionality reduction method is needed. We have considered to apply the Principal Component Analysis with a Whitening process (WPCA) [18] in our framework. This method not only reduces the dimension of the feature vector, making it more compact, but also can be used even on small-sample-size cases [18].

The rest of this paper is organized as follows: in section 2, related work is analyzed; in section 3, the proposed framework is introduced, and the strategy to learn the number of clusters in a region is presented; section 4 shows experimental results of the proposed method in comparison with some related state-of-the-art descriptors; finally, conclusions are given in section 5.

2 Related work

LBP is one of the most popular face image descriptors [19,20]. It was introduced in this area in 2004, motivated by the fact that faces can be seen as a composition of micro-patterns which can be well described by this operator [9]. The original LBP [21] describes these local texture patterns by thresholding the comparison results between

the intensity value of the center pixel and its 3×3 neighborhood. The resulted binary values are then concatenated together and encoded as an integer. This encoding process is illustrated on Figure 1. The operator is invariant to grayscale monotonic variations since it only takes into account if the surrounding pixels values are brighter or darker than the center pixel value. The original method was later extended for using a circular neighborhood of different radius sizes and considering different numbers of equally spaced pixels on the defined circle [22]. In the same work [22], it was shown that more than 90% of the texture information (lines, edges, corners) is contained on 58 patterns which have at most two bitwise 0 to 1 or 1 to 0 transitions; so these patterns were called uniform LBP and, in this case, a single label is assigned to all remaining patterns.

In the past few years, a number of variants of the original operator have been proposed for improving different aspects of the method [20]. Some of the extensions aim at enhancing the discriminative capability of the operator, such as the improved LBP (ILBP) [10], in which both the pixels in the circular neighborhood and the center pixel are compared against the mean intensity value of them. Another is the extended LBP (ELBP) [23], which encodes the gradient magnitude image in addition to the original image in order to represent the velocity of local variations. There are some extensions that have been proposed for improving the robustness, such as the local ternary patterns (LTP) [16] which includes a 3-level generalization coding scheme and is more resistance to noise. However, this operator is no longer invariant to monotonic grayscale transformations. There are also some extensions which have concentrated on choosing an appropriate neighborhood for the encoding process (e.g. the number and distribution of the sampling points as well as the shape and size of the neighborhood). One of the examples is the Multi-scale Block LBP (MB-LBP) [24], in which the average intensity values of neighboring rectangular blocks are compared rather than single pixels. This allows to capture macro-structures of face images. Three-Patch LBP and Four-Patch LBP [25] are also patch-based operators, and their experimental results are very promising.

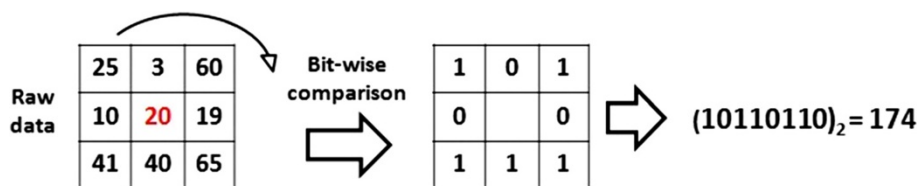


Figure 1 LBP encoding process. This figure describes the LBP encoding process in which each pixel is compared with its eight neighbor pixels and the comparisons are encoded in a binary number representing the LBP code.

Most of the descriptors mentioned above use the original strategy of Ahonen et al. [9] for facial representation. The scheme consists of dividing the face image into rectangular regions, from which local histograms of the extracted local patterns are obtained. Afterwards, the histograms of all regions are concatenated into a single spatially enhanced feature histogram that encodes both the local texture and the global shape of face images. Under this strategy, deciding the number and size of blocks is usually a problem, especially when there are different appearance variations on the face. A finer division usually makes the descriptor more discriminative but sometimes, for example when there are expression variations, will bring some problems. This is illustrated on Figure 2, where it can be appreciated that in the case of expression variations, a finer division can affect the recognition process because small blocks around some face areas, such as mouth and eyes, are shifted to neighbor blocks. Just a few methods on the literature aim at modifying the spatial division strategy. In [26] and [27], many subregions are obtained by shifting and scaling a rectangular region over the face image and boosting is used for selecting the most discriminative regions of different sizes at different positions. Overlapped subregions have also been used [28]; as well as circular [29] and triangular [30] regions.

The spsLBP method [17] is another approach proposed to solve the blocks division problem. It uses a simple clustering method to segment the pixels in a region by considering their intensity values. In spsLBP, the face image is first divided into a few coarse rectangular regions and then the pixels in each region are regrouped by their semantic meanings. Histograms of LBP codes are computed from

the obtained sets and concatenated as in the traditional scheme. It is a very simple idea in which the local patterns are associated with their semantic meaning instead of their spatial position only. This strategy allows to group most of the relevant pixels into corresponding sets even in the presence of some shifting. It should be noticed that, for simplicity, intensity values were used for associating the pixels, but some other attributes such as contrast, luminance, texton, etc., could also be used. It was shown in [17] that this strategy outperforms the traditional regular division. However, more robust LBP variants were not considered in that work. On the other hand, the number of pixel sets for each region was set equal which can be inappropriate in some cases. In Figure 3, it is shown that different face regions can contain a different amount of variable information. By intuition, some regions rich in texture, like areas around eyes, should contain more information; thus, a large number for pixel groups with different semantic meanings should be set while others, like cheek, can be almost homogenous. Hence, we believe both, more robust descriptors and proper number of sets for each region, can boost the performance of this framework.

3 Face feature extraction using semantic pixel set-based local patterns

The most often used strategy for obtaining face descriptors is based on spatial histograms of local patterns. However, the grouping statistical process only considers the spatial information of the pixels, and this may be the reason for non-corresponding sub-blocks matching when there are large expression variations. Hence, we present

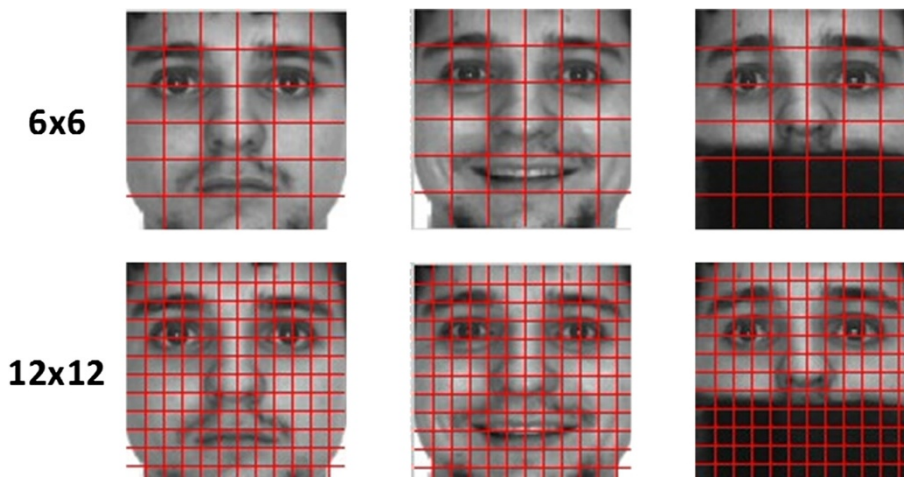
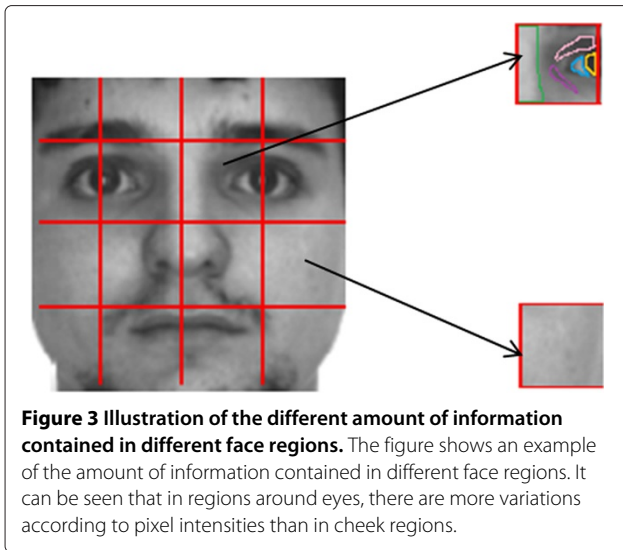


Figure 2 Traditional regular face blocks division. This figure illustrates the effect of using different sizes for regular blocks division. It is shown that when there are occlusions on faces, a finer division allows to have more blocks for comparisons (valid areas); meanwhile, in the case of expression variations, a finer division produces a decrease on the recognition rates because small blocks around some face areas, such as mouth and eyes, are shifted to neighbor blocks.



a strategy for associating local patterns in a face region by their semantic meaning in an adaptive way, in order to exploit better the information within each rectangular face region.

The process of face feature extraction using the general framework of semantic pixel set-based local patterns is illustrated on Figure 4. First, a face image is divided into a few regular blocks of a given size, and the number of pixel sets (N_i) for block i , is learned according to the information entropy of the block. Then, pixels in this block are re-grouped into N_i sets according to their semantic meaning. Once we have different sets of pixels for each block,

histograms of local patterns are extracted from each of them. Finally, all features are concatenated together and enhanced by the WPCA method.

3.1 Learning the number of sets based on the information entropy

The entropy is a term defined in information theory as a measurement of the uncertainty associated with a random variable [31]. It is relevant to the quantity and variability of the information. Here, we assume that the pixel intensity value is a random variable; thus, we can use the histogram of the intensities in each face block to approximate the probability density function (PDF) for computing the information entropy. Applied to our case, the larger the entropy value is, the more information a face block should contain, and thus more clusters should be set.

The entropy value of the face block i can be then defined as

$$S(i) = \sum_{k=1}^n p(x_k) \log_2 \left(\frac{1}{p(x_k)} \right) = - \sum_{k=1}^n p(x_k) \log_2 p(x_k) \quad (1)$$

where $p(x_k)$ is the probability of the pixel x with intensity value k in the histogram of the block.

In our proposal, the entropy following Equation 1, is computed from the intensity histograms of the coarse-divided regions for all face images in the training set. Then, the average entropy value of a block in all images is used as the corresponding regional entropy. Although some images in the training set might be affected by noise, the average entropy values can still reflect the information quantity differences among different facial regions.

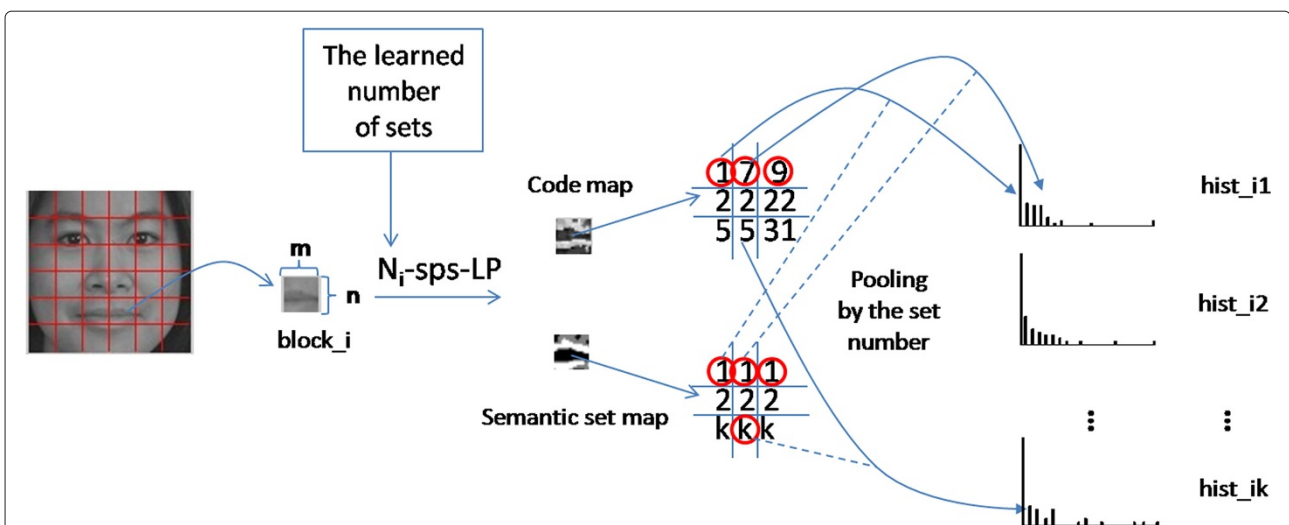


Figure 4 The flowchart of the semantic pixel set-based local patterns using information entropy. This figure describes the complete flowchart of the proposed method. First, a face image is coarsely divided into a few blocks of a given size. With the learned number of pixel sets for each face block according to the information entropy, pixels are re-grouped to different sets according to their semantic meaning. Histograms of local patterns are then extracted from each set. Finally, all features are concatenated together and enhanced by the WPCA method.

Finally, a monotonic transform function is used for mapping the entropy value to the number of sets. The whole process is described in Figure 5.

The monotonic transform function $F(x_i)$ in this paper, is implemented by using a linear function as follows:

$$F(x_i) = (x_i - x_{\min}) / (x_{\max} - x_{\min}) \times (\text{new}_{\max} - \text{new}_{\min}) + \text{new}_{\min}, \quad (2)$$

where x_i is the average entropy for block i , x_{\min} and x_{\max} are the minimum and the maximum entropy values from all regions, new_{\min} is the least sets the region should be divided while new_{\max} is the maximum number of sets that can be obtained in a region. If the output of $F(x_i)$ is not an integer number, it can be rounded to be an integer value.

In this work, we have decided to use $\text{new}_{\min} = 2$ and $\text{new}_{\max} = 8$ and a coarse blocks division of 6×6 , aiming at having a good trade-off between the computational cost and the proper use of the local spatial information.

Illustrated in Figure 6 is the number of sets learned with the proposed algorithm, using the mentioned configuration, for a given training set. In the image, the brightest parts represent those regions with number of sets equal to 8 and the darkest parts correspond to the number of sets 2, while the gray parts represent integers between 2 and 8. It can be seen that those blocks corresponding to the eyes and nose contain more information than the rest of the parts. This corresponds to our intuition and also obeys the traditional weighted maps used for face recognition. Hence, we believe that using different number of sets for each block will enhance the discriminant ability of the method proposed in [17], where a fixed number of sets was used. Besides, it can also help to make better use of the spatial information.

3.2 Semantic pixel sets based local patterns

Once the number of pixel sets (N_i) in each face block is learned on the training phase, the proposed semantic

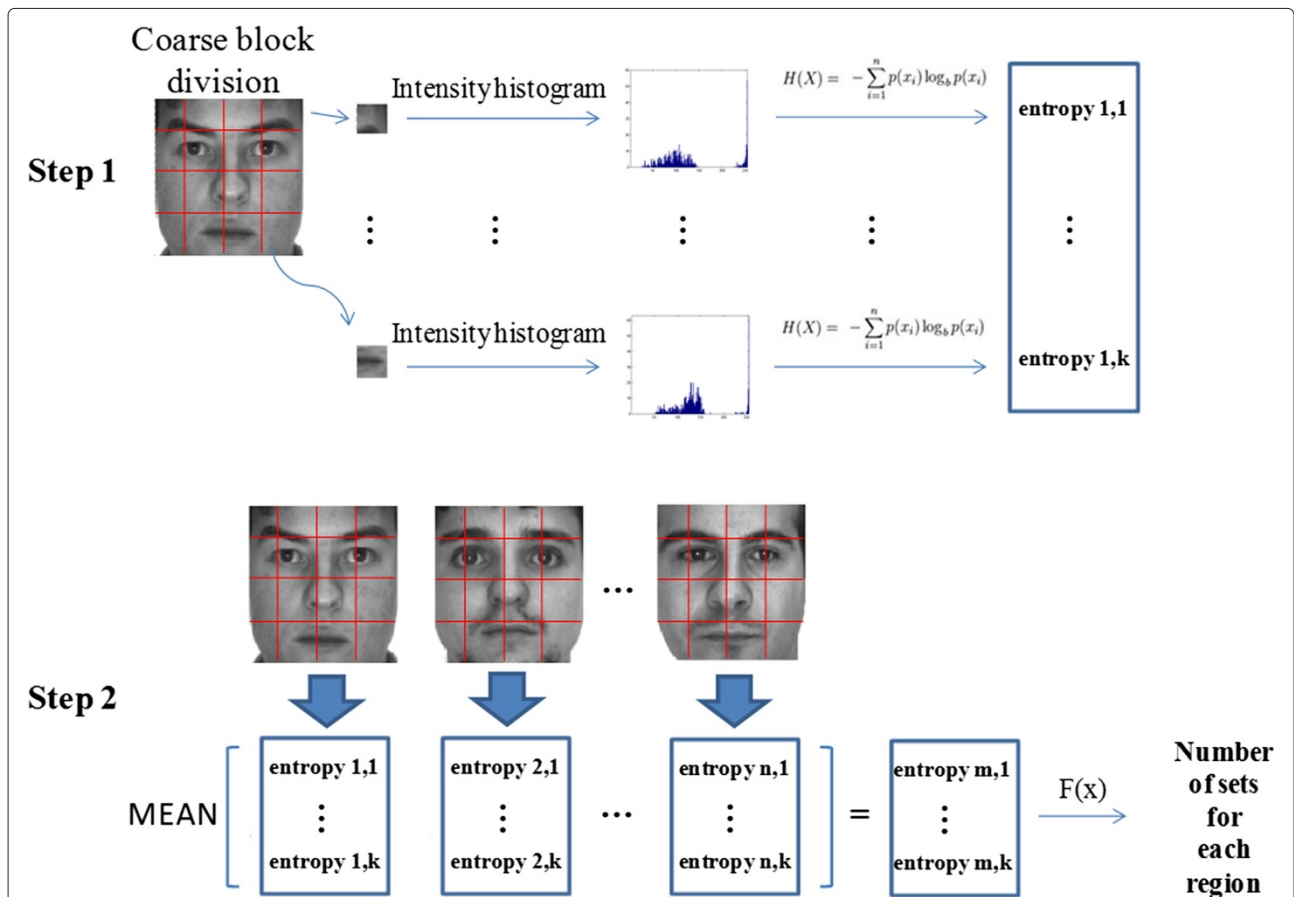


Figure 5 The details for learning the number of sets in each face block using information entropy. This figure describes the process of learning the number of sets for each face block using the information entropy. First, the pixel intensity value in each face block is treated as a random variable; thus, we can use the histogram of the intensities in each face block to approximate the probability density function (PDF) for computing the information entropy. Then, the average entropy value of all face images is used as the regional corresponding entropy. Finally, a linear function is used as a mapping from the entropy value to the number of sets.



Figure 6 The learning results of set number in each face block. In the figure, the number of sets learned with the proposed algorithm based on entropy, for a given face division, is shown. It can be seen that those blocks corresponding to the eyes and nose contain more information than the rest of the parts.

pixel set-based strategy for obtaining the histogram features can be used in the recognition process. First, the pixel intensity values on block i are sorted and clustered uniformly in N_i sets, as it is illustrated in Figure 7. Under this strategy, for a fixed number of sets, the division of a block will always be the same although some have monotonic variations; so if the used local pattern operator, such as LBP, is invariant to monotonic grayscale variations, the final descriptor will also inherit this property (Additional file 1).

As was mentioned before, not only the LBP operator but also any other local pattern-based encoding method can be used with our strategy since the final representation will be given by the histograms of codes computed from each pixel set. So, we will have for each coarse block the corresponding semantic set map and the codes map computed by the encoding method (e.g. LBP, LTP, TPLBP, etc.).

Using both, the semantic set map and the computed codes map, the histogram for the set $S(i, n)$, with $n \in [1, \dots, N_i]$, can be obtained by

$$H_{i,n}(l) = \sum_{x,y \in S(i,n)} I\{\text{codes_map}(x,y) = l\}, l = 0, 1, \dots, m-1 \quad (3)$$

where $\text{codes_map}(x,y)$ is the local pattern obtained at position (x,y) and m is the number of code labels.

Finally, all feature vectors from all sets of all blocks ($1 : t$) will be concatenated together to represent a face image:

$$X = [H_{1,1} H_{1,2} \dots H_{1,N_1} \dots H_{t,N_t}]. \quad (4)$$

In the following, we will call this face image descriptor, semantic pixel set-based local patterns using information entropy (en-spsLP). As have been explained, the local patterns (LP) can be any histogram descriptor based on a local operator such as LBP and its different extensions.

3.3 Dimensionality reduction using WPCA

In order to take more advantage of the spatial information, overlapping regions can be used to extract the en-spsLP features. However, this increases the total dimension of the feature vector and can bring the curse of dimensionality. Hence, a feature reduction method should be applied in this case, in order to get a more compact representation.

There are different methods in the literature for dimensionality reduction. Most of the supervised methods usually applied in face recognition like LDA, although have shown good results, require more than two images per person for training, which cannot be always satisfied in real applications. It was proposed in [18] to apply Principal Component Analysis with a Whitening process (WPCA) to solve the so-called 'Single Sample per Person' problem. This method has been recently used with different face descriptors such as Local Gabor Binary Patterns [32] and POEM [33], showing a very good performance even when only one or a few images are available per person. For those reasons, we decided to apply WPCA for the dimensionality reduction in our framework.

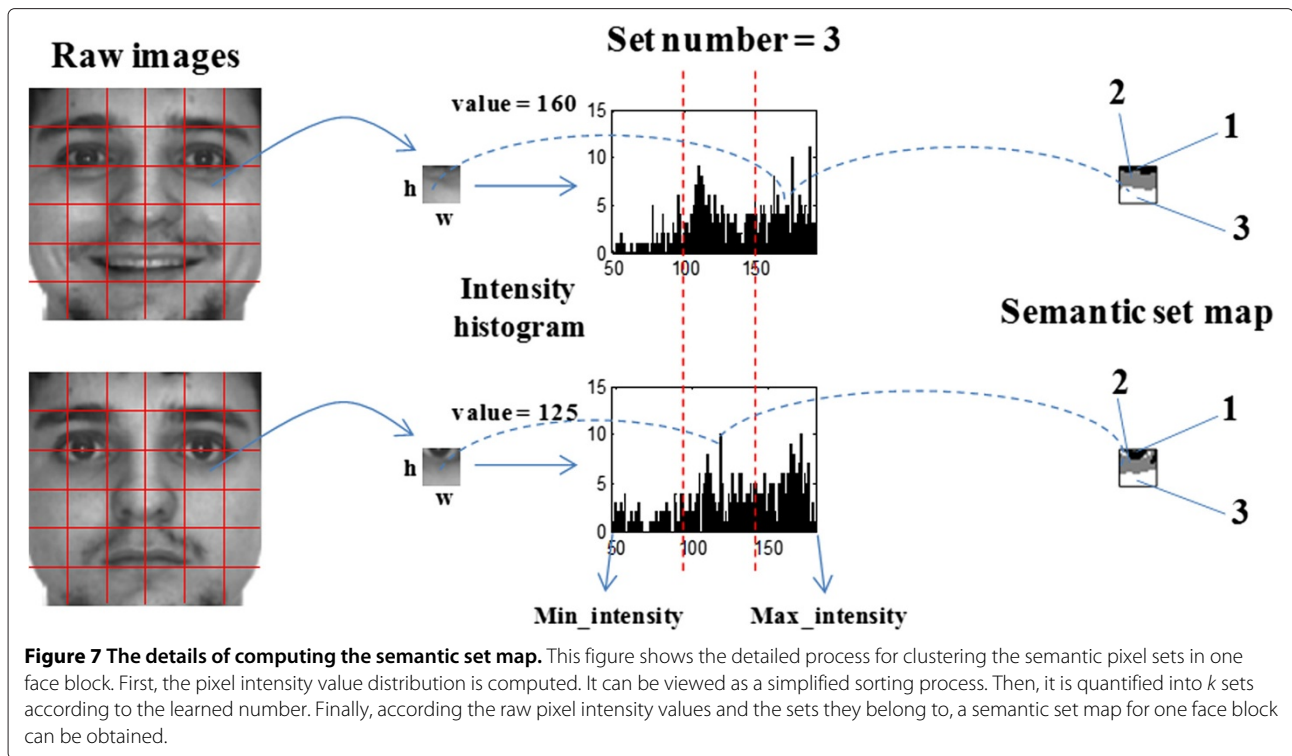
Under this method, after a feature vector, X , is projected into the lower dimensional feature space found by PCA, $u = W_{\text{PCA}}X$, it is normalized with a whitening transformation:

$$w = \Lambda_M^{-1/2} u, \quad (5)$$

where $\Lambda_M^{-1/2} = \text{diag}\{\lambda_1^{-1/2}, \lambda_2^{-1/2}, \dots, \lambda_M^{-1/2}\}$ and λ_i are the eigenvalues of the covariance matrix. This process aims at reducing the negative influences of the leading eigenvectors, as well as magnifies the discriminating details encoded in the trailing ones.

4 Experimental evaluation

Verification and identification experiments were conducted in order to evaluate the performance of the proposed method. Two popular databases: FERET [34] and AR [35], were used for identification experiments, while the LFW [36] database was used for verification. In those experiments where WPCA was not applied, the χ^2 distance was used to compare the obtained descriptors from face images, otherwise cosine distance will be used. In the case of identification, the nearest neighbor classifier



was applied, and the top-rank recognition rate was used to measure the performance of the methods. In the case of verification on LFW, we have followed the evaluation protocol, and the estimated mean classification accuracy with the standard error ($\hat{\mu} \pm SE$) was used for the evaluation. All images were photometric normalized with the preprocessed sequence proposed by Tan and Triggs [16].

The two databases used for identification are composed by images captured under controlled environments. The FERET database [34] contains images with a lot of variations in expression, lighting and aging, divided into five subsets: Fa (gallery set), composed by frontal images of 1,196 subjects; Fb containing 1,195 face images with variations in expression; Fc subset, which contains 194 images with variations in lighting; Dup-I with 722 face images taken with an elapsed time with respect to the images in the gallery set; and Dup-II, a subset of Dup I, which contains 234 images in which the elapsed time is at least 1 year. On the other hand, the AR database [35] was created to test face recognition methods in front of various expressions, different illuminations and occlusions. It contains more than 3,200 face images of 126 people captured on two different sessions. Each person has up to 13 images per session. We randomly selected 100 different subjects (50 males and 50 females) and the neutral expression image of every person in each session was used as gallery and the rest of them with different expressions, lighting and occlusions were used for testing. Images from both

databases were cropped to 114×114 . Thus, using a coarse block division of 6×6 , the blocks size will be 19×19 .

Different from the former two databases, the LFW [36] database contains 13,233 images that were obtained under different unconstrained environments. The images are from 5,749 different individuals, and 1,680 of them have two or more images. In our experiment, we follow the standard training and testing protocol. We have used here the 'View 1' for learning the number of sets for each face block, and the 'View 2' for the final testing. Under this protocol, 6,000 pairs of images are compared in the evaluation; the half of them correspond to images from the same person and the other half not. The testing data are divided into 10 evenly distributed sets and the test is repeated 10 times, using one set for testing and the others for training. It should be noted that our proposal was tested with the original data, without correcting the few labeling errors in the database. Besides, the aligned version (provided by [37]) of face images was used, and all of them were cropped to 126×110 . In this case, overlapped coarse blocks of 18×22 were used.

4.1 The contribution of semantic pixel sets to different LBP based descriptors

The aim of the first experiment is to show that the semantic pixel set (sps)-based strategy, makes not only the original LBP but also other LBP-based descriptors, more robust and stable under different variations of facial

appearance. It is expected that by using more robust descriptors, better results can be achieved. In order to make a fair comparison with [17], we use in this experiment the same fixed number of sets for each region, i.e. 6 or 8 sets. Besides, the original uniform LBP, the Local Ternary Patterns (LTP) [16] and the Three-Patch LBP (TP-LBP) [25] with the coarse initial division are tested. The obtained results on the FERET database are listed on Table 1. It can be seen that almost on all cases, the sps outperforms the traditional regular division. Moreover, the more robust the descriptor, the better the results achieved with our proposal. In general, the best performing descriptor is LTP.

In order to have an in-depth comparison between the sps strategy and the traditional block division, we test the LTP descriptor, with each strategy during histogram estimation. In order to make the comparison fair, we do not use the illumination preprocessing in this experiment. Given a coarse blocks division of 6×6 , we compare the results of using LTP directly over this division (LTP), dividing those coarse blocks into more n sub-blocks (n -blockLTP) and dividing them into n sets according to our proposal (n -spsLTP). The results obtained in each subset of AR database with different kinds of variations are listed on Table 2.

As was explained in the Introduction, a finer regular block division is good for some cases (e.g. occlusions) but degrades for some others, especially for expression variations. It can be seen from Table 2, that although the blockLTP presents slightly better results than our proposal for occlusion variations, in the case of expressions, the performance drops off by a significant margin, even worse than the original LTP. In general, our sps strategy has a more stable behavior. In any case, we have to admit the disadvantages of our method in the case of unpredicted occlusions. For instance, the pixel intensities of the original cheek region occluded by sunglasses will not be as dark as the black sunglasses. This will definitely change

Table 1 Top rank recognition rates on the FERET database

Method	Fb	Fc	Dupl	Dupll	Average
uLBP	91.96	93.29	58.86	49.14	77.61
6-spsLBP	95.73	95.36	69.52	62.39	84.30
8-spsLBP	95.56	95.36	68.28	60.68	83.66
LTP	95.73	96.90	69.39	61.96	84.34
6-spsLTP	97.07	96.90	72.71	67.94	86.65
8-spsLTP	97.15	96.90	71.88	65.38	86.18
TPLBP	91.54	88.14	65.51	55.55	79.65
6-spsTPLBP	95.73	92.26	71.32	64.10	84.77
8-spsTPLBP	95.56	91.23	70.08	62.39	84.05

The contribution of semantic pixel sets to different LBP-based descriptors.

Table 2 Top rank recognition rates on the AR database

Method	Expression	Lighting	Scarf	Sunglasses	Average
LTP	88.66	94.66	74.00	64.83	79.70
6-blockLTP	83.50	99.66	93.83	91.16	92.04
8-blockLTP	82.83	99.33	92.66	89.33	91.04
6-spsLTP	90.00	98.66	91.33	88.83	91.83
8-spsLTP	90.66	99.33	92.33	86.66	92.63

The advantage and disadvantage between sps and regular blocks division.

the sorting results of the pixel intensity values. The pooling process will thus go wrong. So, in the future, we will try some appearance-based methods for clustering in order to get a more stable block division result.

4.2 The contribution of entropy-based learning algorithm for estimation the number of sets in each block

The aim of the this experiment is to demonstrate the contribution of using the information entropy for learning the number of sets for each face block. In this case, we compare the results of the same three descriptors with a fixed number of sets for each block (best results from Table 1) and using information entropy for learning the number of sets (en-sps). The obtained results are shown on Table 3. It can be appreciated that by using different number of sets for each region, better results are achieved in almost all cases. Besides, it can be said that the proposed learning method is useful for deciding the number of sets in each face block.

4.3 Face recognition using information entropy based spsLTP

In order to further exploit the capabilities of the proposed descriptor and to make better use of the spatial information, in this experiment we use the overlapping regions to derive the face features. Since in both experiments above, the LTP-based descriptor performs the best, we are going to use it in this experiment. In this case, keep the same blocks size but with five pixels of overlapping between neighbor blocks. When more blocks are involved, the final

Table 3 Top rank recognition rates on the FERET database

Method	Fb	Fc	Dupl	Dupll	Average
6-spsLBP	95.73	95.36	69.52	62.39	84.30
en-spsLBP	96.31	96.39	71.60	67.52	85.84
6-spsLTP	97.07	96.90	72.71	67.94	86.65
en-spsLTP	97.23	96.39	74.93	70.94	87.67
6-spsTPLBP	95.73	92.26	71.32	64.10	84.77
en-spsTPLBP	97.74	96.90	72.02	65.38	86.52

The contribution of entropy based learning algorithm for the estimation of the number of sets in each face block.

feature vector size becomes many times the original one. Hence, the use of a feature reduction method is needed. As it was explained above, the WPCA method is applied in this case. The obtained results, compared with some other face descriptors on FERET database are shown on Table 4. It can be seen that for spsLTP, better results can be achieved with the overlapping version (spsLTP-ov). Besides, when the learned number of sets for each block is used, the results can get further improvement, compared with the results obtained by using a fixed number of sets for all blocks. Moreover, the benefits of using WPCA are demonstrated. In this database, when using overlapping blocks, we have found a total of 1,298 sets. This means that the descriptor (en-set-spsLTP-ov) has a dimension of 153,164 ($1,298 \times 59 \times 2$). We have selected only 850 features by using WPCA (en-set-spsLTP-ov-WPCA) and a better result is obtained. So, by applying WPCA, we get a more compact and discriminative descriptor.

It can also be seen on the table that our results are comparable with some of the state-of-the-art methods such as the Local Gabor Binary Patterns Histograms Sequence (LGBPHS) [32], the Learned Local Gabor Patterns (LLGP) [39], the Histograms of Gabor Ordinal Measures (HOGOM) [40], the Patterns of Oriented Edge Magnitudes (POEM) [42] and Discriminative Local Binary Patterns (DLBP) [41].

4.4 Descriptors comparison in unconstrained environment

The aim of this last experiment is to show the effectiveness of our proposal in a more challenging face database, the LFW, for the uncontrolled face verification task. Since our method is unsupervised, only related works are compared. All images of the aligned version [36] are cropped

to be 126×110 around the center. Blocks of 18×22 pixels are used to extract the en-spsLTP features, overlapped by six pixels in the rows and eight pixels in the columns. During model selection, the training set in 'View 1' was used to learn the number of sets for each block by using the information entropy. For testing in 'View 2', the training set is used to train the WPCA axes and find the best threshold for determining if a comparison corresponds to the same person or not. In Table 5, we compare our method with other descriptors evaluated in [37] and [43] under the same protocol. We can find that in the list of unsupervised methods, our proposed method is comparable with the other descriptors. Since this dataset is very challenging, we believe our method can achieve a better result by using a more complex nonlinear classifier learned from a supervised way.

5 Conclusions

This paper proposes a face representation framework called histogram of semantic pixel set-based local patterns using information entropy (en-spsLP). First, the number of pixel sets is learned according to the information entropy in each face block. Then, during the histogram estimation, the code of the local pattern is pooled according to the original pixel intensity distribution. Finally, all the histograms are concatenated together and enhanced by the WPCA. The proposed method is easy to implement and the speed of the feature extraction is very fast. At the same time, the results are comparable with some of the state-of-the-art methods. Future work is to try other clustering criterion (e.g. by texton) in order to achieve more robustness to unpredicted occlusions. Besides, the optimal values for initial block size and position based on face landmarks can be further analyzed.

Table 4 Top rank recognition rates on the FERET database

Method	Fb	Fc	Dupl	Dupll
LGBPHS [32]	94.00	97.00	68.00	53.00
LGBPWP [18]	98.10	98.90	83.80	81.60
POEM [33]	97.60	96.00	77.80	76.50
POEM + WPCA [38]	99.60	99.48	88.78	85.00
LLGP [39]	99.00	99.00	80.00	78.00
HOGOM [40]	98.10	99.50	83.60	82.00
DLBP [41]	99.00	99.00	86.00	85.00
6-spsLTP-ov	97.65	99.48	76.73	72.22
en-spsLTP-ov	98.41	99.48	79.36	76.92
6-spsLTP-ov + WPCA	99.41	99.48	84.90	81.62
en-spsLTP-ov + WPCA	99.74	99.48	88.78	85.89

Face recognition using entropy based spsLTP-ov + WPCA in controlled environment.

Table 5 Mean scores on the LFW database on 'image-restricted configuration' and aligned images

Method	Performance
uLBP [37]	0.6824
Gabor (C1) [37]	0.6849
TPLBP [37]	0.6926
FPLBP [37]	0.6818
SIFT [37]	0.6986
V1-like [43]	0.6421 ± 0.0069
V1-like+ [43]	0.6808 ± 0.0044
LARK [44]	0.7223 ± 0.0049
POEM [33]	0.7520 ± 0.0073
POEM + WPCA [45]	0.8113 ± 0.0053
en-spsLTP-ov + WPCA	0.8050 ± 0.0033

Face recognition using entropy based spsLTP-ov + WPCA in uncontrolled environment.

Additional file

Additional file 1: Proof of the monotonic grayscale invariant property of the semantic pixel set strategy.

Competing interests

The authors declare that they have no competing interests.

Acknowledgement

This work is funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA06030300), National Basic Research Program of China (Grant No. 2012CB316300), National Natural Science Foundation of China (Grant No. 61075024, 61273272, 61103155) and International S&T Cooperation Program of China (Grant No.2010DFB14110).

Author details

¹National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science, P.O. Box 2728, Beijing 100190, People's Republic of China.

²Advanced Technologies Application Center, 7th Avenue #21812 b/ 218 and 222, Havana, Cuba.

Received: 1 November 2012 Accepted: 15 April 2014

Published: 6 May 2014

References

1. AK Jain, SZ Li, *Handbook of Face Recognition*. (Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005)
2. W Zhao, R Chellappa, PJ Phillips, A Rosenfeld, Face recognition: a literature survey. *ACM Comput. Surv.* **35**(4), 399–458 (2003)
3. B Heisele, P Ho, J Wu, T Poggio, Face recognition: component-based versus global approaches. *Comput. Vis. Image Underst.* **91**(1–2), 6–21 (2003)
4. X Tan, S Chen, Z Zhou, F Zhang, Face recognition from a single image per person: a survey. *Pattern Recognit.* **39**(9), 1725–1745 (2006)
5. R He, BG Hu, WS Zheng, XW Kong, Robust principal component analysis based on maximum correntropy criterion. *IEEE Trans. Image Process.* **20**(6), 1485–1494 (2011)
6. Á Serrano, IM de Diego, C Conde, E Cabello, Recent advances in face biometrics with Gabor wavelets: a review. *Pattern Recognit. Lett.* **31**(5), 372–381 (2010)
7. Z Lei, S Liao, M Pietikäinen, SZ Li, Face recognition by exploring information jointly in space, scale and orientation. *IEEE Trans. Image Process.* **20**, 247–256 (2011)
8. M Pietikäinen, A Hadid, A Zhao, T Ahonen, *Computer Vision Using Local Binary Patterns*. (Springer-Verlag, London Ltd, 2011)
9. T Ahonen, A Hadid, M Pietikäinen, Face recognition with local binary patterns, in *European Conference on Computer Vision (ECCV)*. Prague, Czech Republic, 11–14 May 2004, pp. 469–481
10. H Jin, Q Liu, H Lu, X Tong, Face detection using improved LBP under Bayesian framework, in *International Conference on Image and Graphics (ICIG)*. Hong Kong, 18–20 Dec 2004, pp. 306–309
11. S Liao, ACS Chung, Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude, in *Asian Conference on Computer Vision (ACCV)*. Tokyo, 18–22 Nov 2007, pp. 672–679
12. S Liao, X Zhu, Z Lei, L Zhang, SZ Li, Learning multi-scale block local binary patterns for face recognition, in *International Conference on Biometrics (ICB)*. Seoul, 27–29 Aug 2007, pp. 828–837
13. Z Guo, L Zhang, D Zhang, X Mou, Hierarchical multiscale LBP for face and palmprint recognition, in *International Conference on Image Processing (ICIP)*. Hong Kong, 26–29 Sept 2010, pp. 4521–4524
14. S Liao, M Law, A Chung, Dominant local binary patterns for texture classification. *IEEE Trans. Image Process.* **18**(5), 1107–1118 (2009)
15. Z Guo, L Zhang, D Zhang, A completed modeling of local binary pattern operator for texture classification. *IEEE Trans. Image Process.* **19**(6), 1657–1663 (2010)
16. X Tan, B Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process.* **19**(6), 1635–1650 (2010)
17. Z Chai, H Mendez, R He, Z Sun, T Tan, Semantic pixel sets based local binary patterns for face recognition, in *accepted in Asian Conference on Computer Vision (ACCV)*. Daejeon, 5–9 Nov 2012
18. W Deng, J Hu, J Guo, Gabor-eigen-whiten-cosine: a robust scheme for face recognition, in *AMFG*. Beijing, 16 Oct 2005, pp. 336–349
19. S Marcel, Y Rodriguez, G Heusch, On the recent use of local binary patterns for face authentication. *Tech. Rep.*, 06–34, Idiap, 2006
20. D Huang, C Shan, M Ardabilian, Y Wang, L Chen, Local binary patterns and its application to facial image analysis: a survey. *IEEE Trans. Syst. Man Cybernetics-Part C: Appl. Rev.* **41**(6), 765–781 (2011)
21. T Ojala, M Pietikäinen, D Harwood, A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* **29**, 51–59 (1996)
22. T Ojala, M Pietikäinen, Mäenpää T, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
23. X Huang, SZ Li, Y Wang, Shape localization based on statistical method using extended local binary pattern, in *International Conference on Image and Graphics (ICIG)*, (2004), pp. 184–187
24. S Liao, X Zhu, Z Lei, L Zhang, SZ Li, Learning multi-scale block local binary patterns for face recognition, in *International Conference on Biometrics (ICB)*. Seoul, 27–29 Aug 2007, pp. 828–837
25. L Wolf, T Hassner, Y Taigman, Descriptor based methods in the wild, in *Faces in Real-Life Images workshop at the European Conference on Computer Vision (ECCV)*. Marseille, 12–18 Oct 2008
26. G Zhang, X Huang, S Li, Y Wang, X Wu, Boosting local binary pattern (LBP)-based face recognition, in *Advances in Biometric Person Authentication, Volume 3338 of Lecture Notes in Computer Science*, ed. by S Li, J Lai, T Tan, G Feng, and Wang Y (Springer Berlin Heidelberg, 2005), pp. 179–186
27. C Shan, S Gong, Owan McP, Conditional mutual information based boosting for facial expression recognition, in *Proceedings of British Machine Vision Conference*. Oxford, UK, Sept 2005
28. T Gritti, C Shan, V Jeanne, R Braspenning, Local features based facial expression recognition with face registration errors, in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*. Amsterdam, The Netherlands, 17–19 Sept 2008
29. T Ahonen, A Hadid, M Pietikäinen, Face description with local binary patterns: Application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(12), 2037–2041 (2006)
30. H Mendez-Vazquez, E Garcia-Reyes, Y Condes-Molleda, A new image division for LBP method to improve face recognition under varying lighting conditions, in *Proceedings of International Conference on Pattern Recognition*. Tampa, Florida, 8–11 Dec 2008, pp. 1–4
31. TM Cover, JA Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. (Wiley, New York, 2006)
32. W Zhang, S Shan, W Gao, X Chen, H Zhang, Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition, in *International Conference on Computer Vision (ICCV)*. Beijing, 17–20 Oct 2005, pp. 786–791
33. NS Vu, A Caplier, Face recognition with patterns of oriented edge magnitudes, in *European Conference on Computer Vision (ECCV)*. Heraklion, Crete, 5–11 Sept 2010, pp. 313–326
34. JP Phillips, H Moon, SA Rizvi, PJ Rauss, The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(10), 1090–1104 (2000)
35. A Martínez, R Benavente, The AR face database. *Tech. Rep.* #24, CVC 1998
36. GB Huang, M Ramesh, T Berg, E Learned-Miller, Labeled faces in the wild: a database for studying face recognition in unconstrained environments. *Tech. Rep.*, 07–49, (University of Massachusetts, Amherst, 2007)
37. L Wolf, T Hassner, Y Taigman, Similarity scores based on background samples, in *Asian Conference on Computer Vision (ACCV)*. Xi'an, China, 23–27 Sept 2009, pp. 88–97
38. NS Vu, A Caplier, Enhanced patterns of oriented edge magnitudes for face recognition and image matching. *IEEE Trans. IP.* **21**(3), 1352–1365 (2012)
39. S Xie, S Shan, X Chen, X Meng, W Gao, Learned local Gabor patterns for face representation and recognition. *Signal Process.* **89**(12), 2333–2344 (2009)
40. Z Chai, R He, Z Sun, T Tan, H Mendez-Vazquez, Histograms of Gabor ordinal measures for face representation and recognition, in *IAPR International Conference on Biometrics (ICB)*,. New Delhi, India, 29 March–1 April 2012, pp. 52–58

41. D Maturana, D Mery, A Soto, Learning discriminative local binary patterns for face recognition, in *International Conference on Automatic Face and Gesture Recognition (FG)*. Santa Barbara, CA, 21–25 March 2011, pp. 470–475
42. NS Vu, HM Dee, A Caplier, Face recognition using the POEM descriptor. *Pattern Recogn.* **45**(7), 2478–2488 (2012)
43. N Pinto, JJ DiCarlo, DD Cox, Establishing good benchmarks and baselines for face recognition, in *Faces in Real-Life Images workshop at the European Conference on Computer Vision (ECCV)*. Marseille, 12–18 Oct 2008
44. HJ Seo, P Milanfar, Face verification using the LARK representation. *IEEE Trans. Inform. Forensics Secur. (TIFS)*. **6**(4), 1275–1286 (2011)
45. NS Vu, A Caplier, Enhanced patterns of oriented edge magnitudes for face recognition and image matching. *IEEE Trans. Image Process.* **21**(3), 1352–1365 (2012)

doi:10.1186/1687-5281-2014-26

Cite this article as: Chai et al.: Explore semantic pixel sets based local patterns with information entropy for face recognition. *EURASIP Journal on Image and Video Processing* 2014 **2014**:26.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com