Abstract—Great progress has been achieved in face recognition in the last three decades. However, it is still challenging to characterize the identity related features in face images. This paper proposes a novel facial feature extraction method named Gabor Ordinal Measures (GOM), which integrates the distinctiveness of Gabor features and the robustness of ordinal measures as a promising solution to jointly handle inter-person similarity and intra-person variations in face images. In the proposal, different kinds of ordinal measures are derived from magnitude, phase, real and imaginary components of Gabor images, respectively, and then are jointly encoded as visual primitives in local regions. The statistical distributions of these visual primitives in face image blocks are concatenated into a feature vector and linear discriminant analysis is further used to obtain a compact and discriminative feature representation. Finally, a two-stage cascade learning method and a greedy block selection method are used to train a strong classifier for face recognition. Extensive experiments on publicly available face image databases such as FERET, AR and large scale FRGC v2.0 demonstrate state-of-the-art face recognition performance of GOM.

Index Terms—Biometrics, Face recognition, Feature extraction, Ordinal measures, Gabor filters.

I. INTRODUCTION

A wide range of applications, from public security to personal consumer electronics, have made face recognition one of the most popular research topics in computer vision and pattern recognition during the past 30 years [1]. Although significant improvement on performance has been achieved recently in this field, it is still a challenging task for real life environments due to the large intra-class facial variations, such as pose, illumination, expression, aging and the small inter-class differences [2]. Feature extraction is one of the most important steps in the process of face recognition in order to overcome these problems.

The main approaches used for face feature extraction are usually grouped into two categories [3], namely holistic subspace analysis and local feature description. Holistic subspace analysis aims to represent the global appearance of a human face with its projections on the subspace, which is usually determined by a given optimization criteria, such as PCA, LDA, and ICA [4]. Subspace face analysis methods perform well under controlled conditions but they have some common shortcomings. For instance, a large and representative training set is usually needed for subspace learning [5] and misalignment of face images may cause significant degradation of recognition performance.

Recently, local feature analysis has achieved promising performance in different face recognition tasks [6][7]. This success is attributed to the following reasons: 1) local features can characterize the facial details which are important for personal identification; 2) there are some local features designed with illumination robust property such as Local Binary Patterns (LBP) [8]; 3) statistical histograms of local features are usually used as facial descriptors, being robust against partial variations of local features; 4) local methods provide more flexibility for recognizing faces with partial occlusions. Taking into account that in general, local features are more suitable for face image analysis, the focus of this paper is a better design of local descriptors for face recognition.

Two representative methods of local feature analysis in face biometrics are Gabor wavelets [9] and Local Binary Patterns (LBP) [10]. Gabor wavelets can extract the local features of facial regions on multiple channels of frequencies and orientations. Both magnitude and phase information of Gabor wavelets can be used for face recognition. For example, Local Matching Gabor method (LMG) [11] was developed to encode the Gabor magnitude features. More than 4,000 Gabor jets are used in LMG and the Borda count strategy is employed to measure the similarity between Gabor magnitude features. Even though feature selection approaches can be used to reduce the number of sampling points in LMG [12], an intrinsic problem of LMG is the requirement of hand crafted sampling points. A typical example of Gabor phase features for face recognition is the Histogram of Gabor Phase Patterns (HGPP) [13]. HGPP can achieve a good performance, but the size of its feature template is 90 times that of a face image. In general, a common shortcoming of Gabor wavelets is the high-dimensionality of feature vectors. On the other hand, LBP has been demonstrated as a powerful texture analysis method with successful application in face recognition [14]. The success of LBP comes from the robust binary encoding between the central pixel and its neighboring pixels. It is invariant to any monotonic transformation of intensity values for all image pixels in a local face region. There are a number of extensions of the original LBP operator, such as Multi-scale
Block LBP (MB-LBP) [15], Local Ternary Patterns (LTP) [16], semantic pixel sets based Local Binary Patterns (spsLBP) [17], Local Salient Patterns (LSP) [18] and Patterns of Oriented Edge Magnitudes (POEM) [19]. Besides, there are some methods that combine both Gabor wavelets and LBP features such as the Local Gabor Binary Pattern Histogram Sequence (LGBP-HS) [20], which uses the histogram of LBP on Gabor magnitude images for face recognition, taking advantage of both LBP and Gabor features; the Effective Gabor Volume LBP code on Gabor Magnitude (E-GV-LBP-M) [21] that explores the texture information jointly in multi-channel Gabor magnitude responses such as scale, orientation and spatial sampling points; and the works from Nicolò and Schmid [22] in which both Gabor Magnitude and Phase responses are encoded with two variants of LBP and used for Multispectral Face Recognition.

Although it is a powerful descriptor, LBP codes may be affected by some noise, as it is illustrated in the example in Figure 1. The figure shows one of the shortcomings of the original LBP operator: in order to make use of the spatial information, a LBP code is jointly determined by multiple in-plane pixels based local comparisons. In this way even a flipping of one of the encoded bits will result in a completely different LBP code. Furthermore, in case of encoding Gabor responses with LBP, as LGBP-HS does, the number of erroneous codes will be eight times if derived from one scale and eight orientations. In this case, if only one bit of local comparison is flipped at the same place, the other seven correct local comparison results can be used and then more useful information will be available for face representation. Considering that LBP is a special case of ordinal measures [23], we believe that some other kinds of ordinal filters can be more robust to overcome these problems. Besides, LBP only considers the qualitative relationship between two pixels rather than ordinal measures among multiple image regions.

![Image 1](https://via.placeholder.com/150)

Figure 1: Illustration of the sensitivity of local binary patterns to noise.

Taking into account the advantages of combining both Gabor wavelets and some kinds of ordinal filters, we propose here a novel local feature analysis method that we have named Gabor Ordinal Measures (GOM). Particularly, we propose to use di-lobe and tri-lobe ordinal filters to effectively encode Gabor magnitude and phase images. The basic idea of GOM is to integrate distinctiveness of Gabor features and robustness of this kind of ordinal measures as a promising solution to jointly handle inter-person similarity and intra-person variations in face images. GOM provides a new idea of applying OM on discriminant feature level for face analysis that was first explored in [24] with promising results. In this paper, we extend the proposed features in order to encode different Gabor components with a larger set of filters.

The rest of this paper is organized as follows. Section 2 reviews the ordinal measures based feature representation and introduces the good properties of the extension to Gabor Ordinal Measures. Section 3 presents the technical details of GOM-based face analysis method. The face classification based on GOM is described in Section 4. Extensive experimental results on publicly available face image databases are given in Section 5. Finally, Section 6 concludes this paper. It should be noted that this work focuses on the face image analysis problem; the necessary preprocessing for a complete face recognition system such as face detection and alignment can be found in the literature [1]. So, the term “face image” or “face pattern” used in this paper usually refers to the normalized face region after preprocessing.

II. ORDINAL MEASURES BASED FEATURE REPRESENTATION

Ordinal measure (OM) is defined as the relative ordering information of multiple variables [25]. In visual representation, ordinal measure can be used to encode the information in intensity level or feature level. For example, intensity level ordinal measure may be the qualitative relationship between the average intensity values of two image regions (Figure 2). If region A is brighter than region B, the ordinal measure is “A>B” and we can use one bit feature code “1” to represent such kind of relationship. Otherwise “B<A” and the ordinal code is “0”. Feature level ordinal measure is the qualitative information computed on the image features, e.g. Gabor features. Visual ordinal measures [23] have the advantage of invariance to monotonic illumination changes and robustness against noise. Variations of the parameters of ordinal measures can generate a discriminative feature template for pattern recognition tasks. In our previous work, intensity level ordinal measures have been successfully applied for iris [23], palmprint [26] and face recognition [27], [28].

A straightforward idea of ordinal feature extraction is to compare the weighted average intensity value of dissociated image regions. Such a process can be implemented by ordinal filtering. If the filtering result is positive, the derived ordinal

![Image 2](https://via.placeholder.com/150)

Figure 2: Ordinal measure of visual relationship between two image regions. An arrow points from the darker region to the brighter one. (a) Region A is darker than B, i.e. A<B. (b) Region A is brighter than B, i.e. A>B.
code is “1”, otherwise it is “0”. For the low probability cases “B=A”, the resulting ordinal code can be assigned with either “1” or “0”. An ordinal filter is usually constituted by multiple positive lobes and negative lobes [23].

Some of the existing facial descriptors, such as LBP and Gabor filters, can be regarded as special cases of ordinal measures [23]. However in this paper we propose the use of multi-lobe differential filters (MLDF) for ordinal feature extraction [23], which gives us a richer representation as it will be discussed later.

Some examples of the MLDF, when Gaussian kernels are used, are shown on Figure 3, and they can be expressed as:

\[
MLDF = C_p \sum_{i=1}^{N_p} \frac{1}{\sqrt{2\pi}\delta_i} \exp\left\{-\frac{(X - \omega_i)^2}{2\delta_i^2}\right\} - C_n \sum_{j=1}^{N_n} \frac{1}{\sqrt{2\pi}\delta_j} \exp\left\{-\frac{(X - \omega_j)^2}{2\delta_j^2}\right\}
\]

(1)

where \(\omega\) and \(\delta\) denote the central position and the scale of a 2D Gaussian filter, respectively, \(N_p\) is the number of positive lobes, and \(N_n\) is the number of negative lobes. Constant coefficients \(C_p\) and \(C_n\) are used to keep the balance between positive and negative lobes, i.e., \(C_pN_p = C_nN_n\).

Figure 3: Some typical ordinal filters. (a) Di-lobe ordinal filters. (b) Tri-lobe ordinal filters.

When comparing MLDF with the original LBP operator, we found that it is more robust to uniform noise. This is shown in Figure 4, where the robustness of the tri-lobe ordinal filter to this kind of noise is illustrated. Although some variants of LBP have been proposed to improve the limitation of traditional LBP operator such as MB-LBP [15] and Directional Binary Encoding (DBE) [29], we think the MLDF is a more fundamental image operator for ordinal feature extraction, which can encode ordinal measures of multiple image regions with more flexible parameter configuration. Some of the variable parameters that can be used, such as the number of positive and negative lobes, orientation, scale and location of each lobe, inter-lobe distance and orientation of a pair of lobes are shown on Figure 5 [23].

Compared to Gabor filters, it can be concluded from Figure 3 and Figure 6, that the shape of odd Gabor filter is similar to the di-lobe ordinal filter, and the sum of the coefficients of Gabor filter is zero, which obeys the rule of ordinal filter. A similar relation also exists between the even Gabor filter and the tri-lobe ordinal filter. However, there are some differences between Gabor and ordinal filter. First, Gabor filters have more lobes and the weight for each lobe is inverse proportional to the distance to the filter center. Besides, what is more important, Gabor filters are continuous, so the boundaries of the neighbor lobes are connected with each other; while ordinal filters are more flexible, in the sense that the boundaries between neighbor lobes can be discontinuous and the lobe size can be set at any positive value. In this way, the lobes of ordinal filters will be more focused on local areas with variable inter-lobe distances and they can better represent the facial structures. As conclusion, Gabor filters are suitable to capture the local information while ordinal filters can encode the relationship between local regions, being both useful in the face recognition task.

Figure 4: Robustness of tri-lobe ordinal measure.

Ordinal measures from image intensity values have shown some robustness on face recognition. However, different from other texture rich biometrics modalities such as iris pattern, it is difficult to find enough robust and identity related ordinal relationships in a face image. Firstly, local intensity variations are not significant in face region because of the similar reflection ratio of facial skin. Therefore the number of building blocks for a reliable ordinal template is limited in intensity level. Secondly, the discriminating power of image intensity ordinal measures is limited because of the similarity between inter-class facial appearance. A good solution to this problem is to derive ordinal measures from feature level, after encoding
some of the face information. A number of works on exploring this idea using different kinds of ordinal filters, have shown to obtain more discriminative features [20][21][22][24].

Similar to previous works, we propose here the use of Gabor filters first in order to enhance the local details of face texture. The benefit is twofold: on the one hand, local variations of Gabor images are significant so that we can derive more stable ordinal measures from them; on the other hand, multi-channel Gabor filters can selectively obtain the texture information at a given orientation and scale. Hence, the face image is decomposed into multiple spectrum channels and the feature space of ordinal measures is enlarged to discriminate face images with inter-class similarity. The ordinal measures derived from Gabor images are named Gabor Ordinal Measures (GOM).

In order to show the good properties of the fused GOM descriptor, we have performed a comparison of the intra and inter distances of different kinds of features. The advantages of GOM over OM and Gabor filters are visually illustrated in Figure 7. The first two columns illustrate two face images from subject A and their three corresponding facial features: Gabor magnitude response (Gabor), Ordinal Measures (OM) and Gabor Ordinal Measures (GOM). The Gabor filter used for obtaining these images was the one with scale 1 and orientation 0 and a di-lobe OM with 4 pixels of inter-lobe distance and 5 × 5 lobe size. The last three columns illustrate the face and feature information of subject B. It should be pointed out that the first four images in the first row are from two subjects with expression variations while the fifth image is from the subject B with heavy noise (captured with a long time elapse and under different illumination condition).

![Figure 7: Illustration of the advantage of Gabor Ordinal Measures](image)

Besides the visual intuition, by computing Euclidean distance between different images of the same subject, we found that OM has shortcomings in terms of distinctiveness. For example, the distance between the first two images corresponding to the same person is 0.0036 and the same value is obtained when the distance between the second and the third image is computed, so the intra class difference and inter-class difference were the same in this case. On the contrary, Gabor information is more distinctive in discriminating different subjects; when computing the distances between the same pairs for Gabor images, 0.0886 is obtained for intra-class and a higher value, 0.1040 for the inter-class pair. However, in the most difficult case, when the last image (the one with illumination problem) is compared with the other two images of the same person, only GOM can obtain a distance value lower than inter-class values. For example, the third image is more similar to the first one than to the last one in both OM and Gabor feature representations, but not in the case of GOM. For GOM intra-class value is 0.0053 and 0.0054 the inter-class one. Although it is a very small difference it could be important. In the case of Gabor the inter-class difference is 0.1040 while the intra-class is 0.1312, which will lead to a false accept or false reject. In summary, GOM is both robust and distinctive in these examples because it inherits the advantages of both ordinal measures and Gabor wavelets. We argue that GOM provides a good solution for handling both intra-class variations and inter-class similarity of face images.

Based on the above analysis, we draw the following conclusions, which motivate the proposal of a novel face recognition method:

1) Ordinal measures are robust image descriptors for face feature representation; and LBP and Gabor wavelets can be regarded as special cases of ordinal image operators.

2) GOM integrates the robustness advantage of ordinal measures and distinctiveness advantage of Gabor wavelets, providing a promising solution for face image analysis.

3) By combining ordinal measures derived from image intensity level and feature level, a better face recognition performance can be achieved.

Although we have demonstrated that GOM is an advanced solution to face image analysis, how to develop a complete algorithm is still a challenging problem due to the flexibility of its parameters configuration. Besides, the discriminative power of one bit GOM is limited, so it is better to combine a group of features to encode the visual vocabulary of local face regions. Then, histograms of ordinal codes can formulate a discriminative Bag of Words (BoW) model [30] for face feature representation. The following section will present the implementation details of GOM.

III. FACE IMAGE ANALYSIS BASED ON GABOR ORDINAL MEASURES

GOM is a general image descriptor on the ordinal measures derived from Gabor features of face images. However, the design of such a simple idea has a number of variables to consider:

1) There are some tunable parameters in Gabor wavelets, such as spatial location, scale and orientation.

2) The convolution of Gabor wavelets with a face image will generate at least four types of Gabor features, i.e. odd and even Gabor filtering results, magnitude and phase responses of Gabor filtering.

3) The variable parameters in ordinal filters include number of lobes, inter-lobe distance, orientation, scale of each lobe, etc.

4) How to combine multiple ordinal measures obtained from Gabor responses is also an important issue.

5) High-dimensional feature vectors are obtained after Gabor filtering and ordinal feature extraction, so a feature
dimension reduction algorithm is necessary to achieve efficient face analysis.

Investigation and determination of all uncertain variables listed above lead to our algorithm of GOM based face analysis method. The process of GOM feature extraction, described in Figure 8, involves mainly five steps: 1) to apply multi-channel Gabor filters on the input face image; 2) to derive ordinal measures from magnitude, phase, real and imaginary components of Gabor images respectively; 3) to jointly encode multiple ordinal measures in local regions as visual primitives; 4) to concatenate the statistical distributions of these visual primitives in face image blocks into a feature vector; 5) to use Linear Discriminant Analysis in order to obtain a compact and discriminative feature representation. The technical details of the whole process of GOM feature analysis are presented step by step in the following subsections.

A. Gabor filtering

The kernels of 2-D Gabor wavelets are similar to the receptive fields of simple cells in the mammalian visual cortex, exhibiting desirable characteristics of spatial locality and orientation selectivity [31]. Therefore the first step of GOM is to use multi-channel Gabor filters to characterize discriminant texture pattern of face images. In our implementation of GOM,
a family of 2D Gabor filters composed by five frequencies and eight orientations (same parameters as used in [31]) is performed on every pixel of a face image [9], which can be formulated as [32]:

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,\nu}\|^2 z^2 + i z^2}{2\sigma^2}} \left[ e^{i k_{\mu,\nu} z} - e^{-\frac{z^2}{2}} \right]$$  (2)

where $\mu \in \{0,\ldots,7\}$ and $\nu \in \{0,\ldots,4\}$ determine the orientation and scale of the Gabor filters and $z = (x, y)$ represents the spatial position. The wave vector $k_{\mu,\nu} = k_\nu e^{i \theta_\mu}$ has a magnitude $k_\nu = k_{\max}/\lambda^\nu$, where $\lambda$ is the frequency ratio between filters and $\phi_\mu = \pi \mu/8$.

The response of a face image $I(x, y)$ to a Gabor filter $\psi_{\mu,\nu}(z)$ is obtained by the convolution:

$$G_{\mu,\nu}(x, y) = I(x, y) * \psi_{\mu,\nu}(z).$$  (3)

The Gabor wavelet coefficient obtained for a given scale and orientation in Equation (3), is a complex number in the following form [9]:

$$G_{\mu,\nu}(x, y) = A_{\mu,\nu}(x, y) \cdot e^{i \theta(x, y)}$$  (4)

where $A$ and $\theta$ represent the magnitude and the phase respectively. The complex Gabor filtering result can also be represented using real part and imaginary part. Different representations of the complex value of Gabor filtering results (magnitude plus phase, or real plus imaginary) show different usages for face feature analysis. Although there are information redundancy among the four kinds of Gabor features, we believe a smart combination (e.g. by LDA) will provide a richer analysis of face images.

Thus, we can obtain four types of feature images after Gabor filtering for each face image, including: 1) Gabor magnitude feature images; 2) Gabor phase feature images; 3) Real Gabor feature images; 4) Imaginary Gabor feature images. Each type of Gabor feature images includes 40 samples corresponding to 5 scales and 8 orientations of Gabor filtering. Besides, each feature image has the same dimension to the original face image.

B. Ordinal measures on Gabor feature images

The four kinds of Gabor feature images provide a set of rich and discriminant texture patterns for ordinal feature representation. Different ordinal feature extraction methods are applied to these four Gabor feature images by considering their individually specific characteristics. The ordinal measures obtained from Gabor magnitude, phase, real and imaginary parts are briefly named as GOM-m, GOM-p, GOM-r and GOM-i respectively.

In order to capture the robust ordinal features in different directions, we use four di-lobe and four tri-lobe ordinal filters in all, with orientation values $0^\circ$, $45^\circ$, $90^\circ$ and $135^\circ$ as shown in Figure 8. The values for the other parameters such as the inter-lobe distance will be discussed in the experimental part. Thus, there are eight multi-lobe differential filters used for ordinal feature extraction on Gabor Magnitude feature images. The obtained ordinal feature set is named GOM-m.

GOM-m$_i^i_{\mu,\nu}(x, y) = \begin{cases} 1 & \text{if } A_{\mu,\nu}(x, y) * \text{MLDF}^1 > 0 \\ 0 & \text{if } A_{\mu,\nu}(x, y) * \text{MLDF}^1 \leq 0 \end{cases}$  (5)

where $A_{\mu,\nu}(x, y)$ is the Gabor magnitude response at orientation $\mu$ and scale $\nu$ for the pixel $(x, y)$ and MLDF$^i$ is the $i$-th multi-lobe differential filter obtained by Eq. 1. The eight MLDFs used for GOM-m extraction include both di-lobe and tri-lobe ordinal filters with four orientations. Some examples are shown in Figure 3. A typical feature representation of GOM-m, shown in Figure 9 a), demonstrates the richness of informative texture details.

Eight ordinal filters used for GOM-m are also performed on Gabor phase feature images. Hence, the ordinal filtering results indicate local phase difference information between positive and negative lobes covered by Gabor phase regions. It has been found that Gabor magnitude information varies slowly while Gabor phase information is sensitive to local variations [20], [13]. However it was shown in [33] that local Gabor phase difference can contain discriminative information and it is more stable than the original phase of Gabor coefficients. Moreover, face recognition performance using the intersection phase angle between two Gabor feature vectors as a descriptor was found comparable to using Gabor magnitude features [33]. Thus, we adopt a similar strategy for ordinal feature extraction on Gabor phase responses. The Gabor phase difference information obtained by ordinal filtering is quantized as GOM-p:

GOM-p$_i^i_{\mu,\nu}(x, y) = \begin{cases} 1 & \text{if } -0 \leq \theta_{\mu,\nu}(x, y) * \text{MLDF}^i \leq 0.5\pi \\ 0 & \text{if } 0.5\pi < \theta_{\mu,\nu}(x, y) * \text{MLDF}^i \leq 2\pi \end{cases}$  (6)

where $\theta_{\mu,\nu}(x, y)$ are the Gabor phase responses at orientation $\mu$ and scale $\nu$. A typical feature representation of GOM-p, shown in Figure 9 b), demonstrates the local detailed texture information of face images.

Both GOM-m and GOM-p are feature level ordinal measures, while GOM-r and GOM-i are intensity level ordinal measures obtained from Gabor filtering results. As was mentioned before, the Gabor feature extraction, which is the first step when computing GOM features, can be viewed as ordinal measures applied to the raw image, if a threshold is used. Here, the Gabor real and imaginary parts are thresholded by zero, obtaining the ordinal measures over Gabor real and imaginary responses (GOM-r and GOM-i):

GOM-r$_i_{\mu,\nu}(x, y) = \begin{cases} 1 & \text{if } \text{Re}_{\mu,\nu}(x, y) > 0 \\ 0 & \text{if } \text{Re}_{\mu,\nu}(x, y) \leq 0 \end{cases}$  (7)

and

GOM-i$_i_{\mu,\nu}(x, y) = \begin{cases} 1 & \text{if } \text{Im}_{\mu,\nu}(x, y) > 0 \\ 0 & \text{if } \text{Im}_{\mu,\nu}(x, y) \leq 0 \end{cases}$  (8)

Finally, four types of GOM feature images are obtained in this step, including:

1) 320 (5 scales * 8 orientations * 8 MLDFs) GOM-m;
2) 320 (5 scales * 8 orientations * 8 MLDFs) GOM-p;
3) 40 (5 scales * 8 orientations) GOM-r;
4) 40 (5 scales * 8 orientations) GOM-i.

It should be noticed that each GOM feature image is a 2D binary template with the same dimension to the normalized face image.
C. Encoding of Gabor Ordinal Measures

There are 720 GOM feature images altogether, which significantly enlarge the feature space of a face image. It is necessary to combine multiple binary codes in GOM feature images to form a discriminant texture primitive and reduce the dimensionality of GOM feature images as well. Since all four types of GOM feature images have eight orientations, it is a good idea to group all eight Gabor ordinal measures in each pixel for a given scale, MLDF, GOM type into a visual code in the BoW model.

\[
\text{GOM-MAP-} \alpha_{\nu}^{i}(x, y) = \begin{bmatrix} \text{GOM-} \alpha_{0,\nu}^{i}(x, y), \text{GOM-} \alpha_{1,\nu}^{i}(x, y), & \ldots, \text{GOM-} \alpha_{7,\nu}^{i}(x, y) \end{bmatrix}
\]

where \( \alpha \in \{ m : \text{magnitude}, p : \text{phase}, r : \text{real}, i : \text{imaginary} \} \) and GOM-MAP-\( \alpha_{\nu}^{i}(x, y) \) is the texture primitive obtained at position \((x, y)\) for the \( i \)-th ordinal measure at scale \( \nu \).

Since the ordinal codes are binary values and eight orientations are used, we will obtain a binary number of 8 bits, which can be converted to a decimal number (a byte) for representing each pixel at each scale:

\[
\text{GOM-MAP-} \alpha_{\nu}^{i}(x, y) = \sum_{u=0}^{7} (\text{GOM-} \alpha_{u,\nu}^{i}(x, y) \ast 2^{u})
\]

Figure 10 illustrates the encoding process. As can be seen, the encoding results of GOM feature images are 90 gray level images \((5 \times 8 \, \text{GOM-m} + 5 \times 8 \, \text{GOM-p} + 5 \, \text{GOM-r} + 5 \, \text{GOM-i})\) with the same dimension to the normalized face image.

D. Histogram representation of GOM

Each visual element in encoded GOM feature maps has the value range from 0 to 255. This means that there are 256 types of visual primitives. Statistical distribution of texture primitives has become a standard description for texture analysis. Therefore a 256-bin histogram representation can be obtained for each block \( R_{n} \) of GOM feature map.

\[
H_{\nu}^{i}(l) = \sum_{x,y \in R_{n}} I\{GOM-\alpha_{\nu}^{i}(x,y) = l\}, l = 0, 1, \ldots, 255
\]

where \( l \) represents the \( l \)-th grayscale value and \( I\{f\} \in \{0, 1\} \) is a boolean indicator of the condition \( f \). Such a representation is robust against intra-class variations of face images and the location information of face texture pattern is also contained in the spatial layout of blocks.

The histogram of each block can be further reduced to contain only \( B \) bins by partitioning the histograms into uniform parts: \([0, \ldots, 256/B - 1], [256/B, \ldots, 2 \ast 256/B - 1], \ldots, [(B - 1) \ast 256/B, \ldots, B \ast 256/B - 1]\). In this paper, we have joined four grayscale values on every partition in order to get a balance between discriminating power and feature dimensionality for GOM. So, we have \( B = 64 \), which means that we obtain a 64-bin histogram for each block. Each histogram has to be normalized based on L2 norm.

Finally, the histograms of GOM-MAP-m, GOM-MAP-p, GOM-MAP-r and GOM-MAP-i are concatenated together to form a feature vector for each block of GOM feature map. The dimension of GOM histogram for each block is thus \( 5,760 \), i.e. \( 64(\text{bins}) \ast 5(\text{scales}) \ast 8(\text{MLDF}) \ast 2(\text{GOM-MAP-m and GOM-MAP-p}) + 64(\text{bins}) \ast 5(\text{scales}) \ast 2(\text{GOM-MAP-r and GOM-MAP-i})\).

E. Feature dimension reduction based on LDA

It is obvious that different components of the original Gabor features have some redundancy and 5,760 dimensional GOM histogram is over-complete for visual representation of facial features in a block. Therefore it is necessary to reduce the dimensionality of feature vector so that a discriminant and compact representation can be achieved. Linear Discriminant Analysis (LDA) is used for this purpose. LDA is a statistical approach for supervised dimensionality reduction and classification which has been widely used on face recognition [34].
LDA determines a set of projection vectors maximizing the between-class scatter matrix ($S_b$) and minimizing the within-class scatter matrix ($S_w$) in the projected feature space.

**IV. FACE IMAGE CLASSIFICATION**

The objective of face image classification is to measure the dissimilarity between two feature vectors of GOM. A number of issues should be considered in face image classification. Firstly, it is desirable to correct possible intra-class variations of GOM features in feature matching step. Secondly, the effectiveness of GOM varies from block to block for face recognition and it is necessary to select only the most effective GOM feature blocks for feature matching. Such a block selection can improve both accuracy and efficiency of face recognition. Thirdly, a smart matching strategy is needed to handle “easy” and “hard” samples separately in face recognition.

![Figure 11: Representation of a face image at three scales.](image)

Although face detection can coarsely normalize the size of face region based on inter-pupil distance, there still exists some scale difference among face images. We found that such a difference is a main cause of face recognition errors. So the normalized face region is used to generate three samples with different scaling factors (Figure 11) and then GOM is performed on these three images. More templates with different scaling factors are useful to reduce the intra-class variations. However, the feature dimensionality is increased to three times of the original size. Therefore a simple greedy block selection method is used in the training phase of our method in order to select a subset of discriminative blocks with a minimum empirical error. The detailed process is described in Algorithm 1.

**Algorithm 1: GOM block selection**

| Input: | Candidate Block Set GOM = \{ GOM_{1,1}, GOM_{1,2}, ... , GOM_{i,j}, ... , GOM_{S,B} \}, where GOM_{i,j} is the candidate feature of j-th block at scale i, S is the number of face scales and B is the number of blocks at one scale. |
| Output: | Selected Block List (SBL) |

Initialize $SBL = \emptyset$

Repeat K times:

for $i = 1$ to $S$

for $j = 1$ to $B$

$Error(i, j) = EER(SBL + GOM_{i,j})$

end

$[p, q] = \arg \min(Erro)$

if $GOM_{p,q} \notin SBL$ continue

else $SBL = SBL \cup GOM_{p,q}$

end

Return $SBL$

Then, on the testing phase, only a small number of effective GOM feature blocks are selected for making the classification.

The Cosine distance is then used as the dissimilarity measure between a pair of GOM blocks. Finally, the sum rule is adopted to combine the Cosine distance as the overall dissimilarity between two face templates.

It is a grand challenge for a single pattern classification system to handle all face images with large variations. Therefore a more practical solution is to employ a divide and conquer strategy. In this paper, a two-stage cascade classifier is used to recognize “easy” and “hard” face images separately. The first stage of face recognition aims to successfully deal with the “easiest” face images. It is expected that most face image matching results can be confidently classified into genuine or imposter samples. However, there must be some face images with ambiguous matching results. Therefore a second stage of classification is necessary to successfully recognize these “hard” samples whose matching scores lie in the overlapping region of the intra and inter class distribution curves. Taking this into account our method uses Algorithm 1 twice in the training phase. First, the whole training data set is used to learn the most effective GOM features which are robust against the general variations of face images. Then, the misclassified samples are used to select the features used for the second classifier. The idea of two-stage face recognition is shown in Figure 12.

![Figure 12: Flowchart of two-stage cascade classifier learning.](image)

It is interesting to investigate the characteristics of GOM features selected in different stages. First, we found that only one third of common GOM features are shared between Stage 1 and Stage 2. As it is shown in Figure 13 a), most GOM features selected in Stage 1 come from small scale face images, which indicates that coarse face features are well suited to discriminate most face patterns. However, these coarse features can not successfully recognize some “hard” samples. Therefore, some finer image features extracted from large scale face images are needed to improve the robustness of face recognition classifier against large intra-class variations. It can be seen from Figure 13 b) that the selected GOM features in the second stage are more evenly distributed in the three scales.

**V. EXPERIMENTAL EVALUATION**

Three publicly available face image databases are used to evaluate the performance of the proposed GOM face recognition method. Sample face images in these databases are shown in Figure 14.
FERET is a traditional face image database [35]. It contains face images with large variations in expression, lighting and aging. There is a total of five subsets in FERET. The gallery set (Fa) contains frontal images of 1,196 subjects. Fb subset contains 1,195 face images with variations of expression. The Fc subset contains 194 images with variations of lighting. Dup1 has 722 face images taken with an elapsed time with respect to the images in the gallery set. Finally Dup2, a subset of Dup1, contains 234 images in which the elapsed time is at least one year. Sample images in each subset are shown in Figure 14 (a). Besides, FERET has a standard training set, which contains 1,002 images from 429 persons.

AR database aims to test the robustness of face recognition methods against various expressions, illumination changes and occlusions [36]. It contains more than 3,200 face images of 126 subjects captured on two different sessions. Each person has up to 13 images per session. We focus here on the scarf occlusion case to test the robustness of GOM against occlusions, so the neutral expression image of every person in each session is used as gallery and the other images with scarf occlusions are used for testing. Images from 90 different subjects (45 male and 45 female) have been used in previous works [21], we use here a larger set with images from 100 people (50 males and 50 females). Samples of the testing images from one person in both sessions are shown in Figure 14 (b). Since there is no specific training set in this database and in order to show the generalization capability of our method, we use the same training set of FERET database in this case.

FRGC 2.0 is a large scale face image database widely used in the literature [37]. It was designed to present challenging images (illumination, expression, blur and time elapsed) to face recognition systems. The database contains 50,000 2D and 3D face images captured under controlled and uncontrolled conditions (see samples in Figure 14 (c)). Two typical experiments on FRGC 2.0 for 2D-2D still images comparisons, i.e. Experiment 1 and Experiment 4, are implemented to test GOM both in controlled and uncontrolled environments. The standard training set used for both experiments consists of 12,776 face images from 222 subjects including 6,360 controlled images and 6,416 uncontrolled ones. 16,028 controlled images from 466 people are used for testing in Experiment 1. The testing data in Experiment 4 includes all 16,028 controlled face images in Experiment 1 and additional 8,014 uncontrolled images are used to test the robustness of face recognition algorithms.

In preprocessing, all images from FERET and FRGC are geometrically normalized to a 128x160 region of interest (ROI) according to the coordinates of the eyes positions. A different size of ROI was used for experiments on AR, because face recognition rate is significantly affected by the image size for occlusion cases. Hence, it is fair to use the same ROI size (80x88 pixels) used in other works [21]. The Gaussian kernel size of each lobe of the ordinal filter is set 5x5 and the sigma is set to 0.5. The normalized face region is then divided into 16x16 blocks and a GOM histogram of 64 bins is used as the feature vector of each block. All face images are processed by the Preprocessing Sequence (PS) method proposed in [38].

It should be noticed that the whole process is made independently in each database. This means that the training phase is different in each case following the standard protocol of each database. In the case of LDA, which is used on feature extraction to obtain a compact feature vector, it reduces the dimensionality of the GOM feature vector of each block to 150 on FRGC, while it reduces the dimensionality to 260 when using the FERET training set.

Since FRGC is the most difficult database we have used, the three scales strategy and the two-stage cascade classifier is only applied in this case for better representing the face variations. We have selected templates with pupil distance 60, 72, 84. In the case of FERET and AR, only one scale for face image is used, which brings a more efficient face recognition system. Then, nearest-neighbor classifier with Cosine distance is used for classification on these two databases.

A. Results on the FERET database

In this section we will show how different parameters of GOM affect the recognition results. In order to show the complementary attributes among different components of GOM feature, first we test GOM-m feature only and try to find the best inter-lobe distance for ordinal filters. Then the fused results with all GOM features will be given later.

We compare the performance of GOM-m features with different inter-lobe distances on FERET Dup1 and Dup2, which are considered to be the two most difficult test subsets.
on this database. From Figure 15 it can be found that the best results can be achieved when inter-lobe distance is equal to 3. Thus we use it in the following experiments.

![Figure 15: The performance of GOM-m with different inter-lobe distances.](image1)

We also test GOM-m with different normalization methods during the histogram statistics. We find that L2-hys normalization [39] performs better than the original L2 norm. L2-hys applies L2 norm followed by clipping (limiting the maximum value of each histogram bin to a threshold) and re-normalization. The recognition results obtained for GOM-m feature with L2-hys normalization are shown on Figure 16 for different threshold values.

Once we have chosen the inter-lobe distance and the normalization method, we test the GOM-m features derived using each of the 8 ordinal filters from the 5 scales Gabor magnitude images. The GOM-m features are fused one by one in a wrapper way according to the order shown in Figure 8. The obtained results for each subset of the database are shown on Figure 17. As can be seen from the graphics, in all cases the recognition rate increases when more GOM-m features are fused, achieving the best performance when all the eight GOM-m features are combined. This also exhibits the complementary attributes of GOM-m features derived using different kind of ordinal filters.

When we compare the performance of different GOM features on FERET Dup1 and Dup2 (Figure 18), we can find that ordinal measures on Gabor magnitude images play a more important role than the other two kinds of GOM features. But these two components of GOM features also provide complementary information and have positive contributions to accuracy. When all components of GOM features are fused together, the best performance can be achieved. This result also demonstrates the success of our information fusion strategy in GOM.

Finally, we have compared the performance of the fused GOM descriptor with state-of-the-art face recognition methods based on Gabor wavelets and LBP. Most of these reference methods were published during the last 5 years. In addition, the top face recognition performance in FERET based on other method such as Patterns of Dominant Orientations (PDO) [40] is also compared with the proposed GOM method. The top-rank recognition rates of GOM and other methods are shown in Table I.

The experimental results show that the proposed GOM descriptor achieves the highest accuracy in all four subsets of FERET. Since we have not used feature selection and classifier enhancement techniques in FERET, to the best of our knowledge, GOM is the best performing face descriptor on FERET in the literature.

The only misclassified sample in fb subset is actually classified correctly, which is caused by the label error in the original database. On the other hand, most of the misclassifications in dup1 and dup2 are due to the coarse alignment
Table I: Top rank recognition rates on FERET database.

<table>
<thead>
<tr>
<th>Method</th>
<th>Fb</th>
<th>Fc</th>
<th>Dup1</th>
<th>Dup2</th>
</tr>
</thead>
<tbody>
<tr>
<td>weighted LBP [14]</td>
<td>97.0</td>
<td>79.0</td>
<td>66.0</td>
<td>64.0</td>
</tr>
<tr>
<td>OM+boosting [27]</td>
<td>98.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LGBP-HS [20]</td>
<td>98.0</td>
<td>97.0</td>
<td>74.0</td>
<td>71.0</td>
</tr>
<tr>
<td>Weighted HGPP [13]</td>
<td>97.5</td>
<td>99.5</td>
<td>79.5</td>
<td>77.8</td>
</tr>
<tr>
<td>E-GV-LBP-M [21]</td>
<td>98.4</td>
<td>98.9</td>
<td>81.9</td>
<td>81.6</td>
</tr>
<tr>
<td>HOGOM [24]</td>
<td>99.2</td>
<td>99.3</td>
<td>82.7</td>
<td>82.1</td>
</tr>
<tr>
<td>CHG [41]</td>
<td>97.5</td>
<td>98.5</td>
<td>85.6</td>
<td>84.6</td>
</tr>
<tr>
<td>DLBP [42]</td>
<td>99.0</td>
<td>99.0</td>
<td>86.0</td>
<td>85.0</td>
</tr>
<tr>
<td>MLG + EJS [12]</td>
<td>99.8</td>
<td>100</td>
<td>88.0</td>
<td>84.2</td>
</tr>
<tr>
<td>HEC [31]</td>
<td>99.0</td>
<td>99.0</td>
<td>92.0</td>
<td>88.0</td>
</tr>
<tr>
<td>POEM + WPICA [19]</td>
<td>99.6</td>
<td>99.5</td>
<td>88.8</td>
<td>85.0</td>
</tr>
<tr>
<td>Tan&amp;Triggs [43]</td>
<td>98.0</td>
<td>98.0</td>
<td>90.0</td>
<td>85.0</td>
</tr>
<tr>
<td>POEM + PDO [40]</td>
<td>99.7</td>
<td>100</td>
<td>91.7</td>
<td>90.6</td>
</tr>
<tr>
<td>MBC-F [44]</td>
<td>99.7</td>
<td>99.3</td>
<td>93.6</td>
<td>91.5</td>
</tr>
<tr>
<td>G-LQP [45]</td>
<td>99.9</td>
<td>100</td>
<td>93.2</td>
<td>91.0</td>
</tr>
<tr>
<td>S[LGBP + LGXP] [46]</td>
<td>99.0</td>
<td>99.0</td>
<td>94.0</td>
<td>93.0</td>
</tr>
<tr>
<td>GOM</td>
<td>99.9</td>
<td>100</td>
<td>95.7</td>
<td>93.1</td>
</tr>
</tbody>
</table>

method by using only the eye positions. We believe even better performance can be achieved with some advanced face alignment method, e.g. similarity transform with more fiducial points.

B. Results on the AR database

In the AR database, our GOM can achieve perfect results on moderate expression (smile and anger) and lighting tasks, and almost perfect results on scarf occlusion case when using the cropped images provided by the database [36]. When the images are downsampled and cropped to a smaller size (as the protocol used in [21]), our GOM can still get excellent results on expression (smile and anger) and lighting, and only a slightly degraded result on scarf case even though a larger testing size (100 subjects) is used. The obtained results and comparisons with other methods for this case are listed in Table II.

Table II: Top rank recognition rates in scarf subset of AR database.

<table>
<thead>
<tr>
<th>Method</th>
<th>Scarf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gabor Magnitude + LDA</td>
<td>71.4</td>
</tr>
<tr>
<td>LBP [14]</td>
<td>47.0</td>
</tr>
<tr>
<td>Gabor Phase + LDA</td>
<td>55.3</td>
</tr>
<tr>
<td>LGBP-HS [20]</td>
<td>82.5</td>
</tr>
<tr>
<td>GV-LGBP-TOP-M [21]</td>
<td>87.4</td>
</tr>
<tr>
<td>E-GV-LBP-M [21]</td>
<td>82.7</td>
</tr>
<tr>
<td>HOGOM [24]</td>
<td>95.0</td>
</tr>
<tr>
<td>GOM</td>
<td>97.6</td>
</tr>
</tbody>
</table>

C. Results on the FRGC database

GOM is implemented on FRGC Experiment 1 and Experiment 4 to test its performance for large scale 2D face recognition, on both controlled and uncontrolled conditions. The focus of this paper is mainly on facial feature analysis. Therefore a fair comparison of GOM with other state of the art face image descriptors is listed on the upper part of Table III and Table IV, where the performance of face recognition is mainly contributed by facial feature analysis methods rather than sophisticated machine learning techniques such as Adaboost. We keep all parameters the same for both experiments to show the generalization performance of the proposed method. We can see from Table III that GOM can achieve the best results in all settings of Experiment 1, even though GOM is extracted from non-overlapping face blocks. Besides, our face recognition result can be further improved with the simple two stage cascade classifier. However it still can not outperform some leading algorithms on Experiment 4 (Table IV), since feature selection with some advanced machine learning techniques such as Boosting [51] or CMI [52] in a larger feature space (overlapping blocks) will also benefit the result, which is demonstrated in the lower part of Table III and Table IV. Hence, we believe GOM still has much potential to be improved.

VI. CONCLUSIONS AND FUTURE WORK

In this work a novel face image descriptor called Gabor Ordinal Measures (GOM) is proposed. First, Ordinal Measures on Gabor (GOM) magnitude, phase, real and imaginary responses at different orientations and scales are extracted. Then, the ordinal information of multiple orientations is encoded into GOM maps. Finally, spatial histograms are extracted from each of the encoded maps. Besides, the use of LDA makes the representation more compact.

The proposed descriptor achieves state-of-the-art results on three popular benchmark databases captured under both

Table III: Verification rates at FAR - 0.1% on FRGC-Experiment 1 for the controlled condition.

<table>
<thead>
<tr>
<th>Method</th>
<th>ROC I</th>
<th>ROC II</th>
<th>ROC III</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEE Baseline [37]</td>
<td>78.8</td>
<td>75.0</td>
<td>70.9</td>
</tr>
<tr>
<td>GFC [47]</td>
<td>-</td>
<td>-</td>
<td>92.5</td>
</tr>
<tr>
<td>KFA [48]</td>
<td>-</td>
<td>-</td>
<td>92.0</td>
</tr>
<tr>
<td>Hybrid Fourier [49]</td>
<td>95.3</td>
<td>93.8</td>
<td>92.3</td>
</tr>
<tr>
<td>MBC-F [44]</td>
<td>98.5</td>
<td>98.0</td>
<td>97.5</td>
</tr>
<tr>
<td>F[LGBP + LGXP] [46]</td>
<td>98.6</td>
<td>97.9</td>
<td>97.2</td>
</tr>
<tr>
<td>GOM (one stage)</td>
<td>98.6</td>
<td>98.3</td>
<td>98.0</td>
</tr>
<tr>
<td>GOM (two stages)</td>
<td>98.7</td>
<td>98.5</td>
<td>98.2</td>
</tr>
<tr>
<td>MB-LBP + Boosting [15]</td>
<td>98.0</td>
<td>97.0</td>
<td>96.0</td>
</tr>
<tr>
<td>LEC + Wrapper [31]</td>
<td>-</td>
<td>-</td>
<td>97.3</td>
</tr>
<tr>
<td>HEC + Wrapper [31]</td>
<td>-</td>
<td>-</td>
<td>98.0</td>
</tr>
</tbody>
</table>

Table IV: Verification rates at FAR - 0.1% on FRGC-Experiment 4 for the uncontrolled condition.

<table>
<thead>
<tr>
<th>Method</th>
<th>ROC I</th>
<th>ROC II</th>
<th>ROC III</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEE Baseline [37]</td>
<td>16.0</td>
<td>15.1</td>
<td>14.0</td>
</tr>
<tr>
<td>GFC [47]</td>
<td>-</td>
<td>-</td>
<td>59.8</td>
</tr>
<tr>
<td>KFA [48]</td>
<td>-</td>
<td>-</td>
<td>76.0</td>
</tr>
<tr>
<td>DCF [50]</td>
<td>81.6</td>
<td>80.6</td>
<td>79.4</td>
</tr>
<tr>
<td>Hybrid Fourier [49]</td>
<td>77.5</td>
<td>77.9</td>
<td>78.2</td>
</tr>
<tr>
<td>MBC-F [44]</td>
<td>82.1</td>
<td>83.0</td>
<td>83.9</td>
</tr>
<tr>
<td>F[LGBP + LGXP] [46]</td>
<td>82.7</td>
<td>83.2</td>
<td>83.6</td>
</tr>
<tr>
<td>GOM (one stage)</td>
<td>85.6</td>
<td>85.2</td>
<td>84.8</td>
</tr>
<tr>
<td>GOM (two stages)</td>
<td>86.2</td>
<td>85.7</td>
<td>85.2</td>
</tr>
<tr>
<td>E-GV-LBP-M+P + CMI [21]</td>
<td>88.7</td>
<td>89.3</td>
<td>89.9</td>
</tr>
<tr>
<td>LEC + Wrapper [31]</td>
<td>-</td>
<td>-</td>
<td>82.8</td>
</tr>
<tr>
<td>HEC + Wrapper [31]</td>
<td>-</td>
<td>-</td>
<td>89.0</td>
</tr>
</tbody>
</table>
controlled and uncontrolled conditions. However as most of the Gabor based methods, GOM also has the disadvantage that the feature extraction part needs more time than LBP or HOG based methods, but it can be finished off-line in enrollment stage. In face recognition stage, the GOM feature extraction is only implemented once for the query face image, and the computational cost of GOM feature extraction is about 700 ms in Matlab program. The matching process of GOM histograms is as efficient as other BoW based face recognition methods and it is expected to achieve millions of matches per second on a multi-core computer.

The main contribution of this paper is a new insight on the facial feature analysis problem based on ordinal measures on discriminative feature level. Firstly, the relationship between ordinal measures and some state-of-the-art descriptors such as LBP is analyzed. Secondly, ordinal measures extracted on feature level are demonstrated to be more discriminative than those extracted on image level. Thirdly, this paper has shown that fusion of image level ordinal measures (GOM-r and GOM-i) and feature level ordinal measures (GOM-m and GOM-p) can achieve better recognition performance. Fourthly, the potential of huge parameter space of ordinal measures for face recognition has not been well exploited, so it is expected that even higher accuracy of face recognition can be achieved with better setting of the GOM parameters. In the future, advanced machine learning techniques such as Adaboost and sparse representation can be used for optimal feature selection.

In summary, the idea of Gabor Ordinal Measures provides a promising solution for face image analysis. The discriminating of Gabor wavelets and robustness of ordinal measures are integrated to handle both inter-person similarity and intra-person variations in face images. Our future work will focus on improving GOM performance by introducing machine learning based feature selection and advanced feature matching strategies.

VII. ACKNOWLEDGEMENTS

The authors are grateful to the anonymous reviewers for their constructive comments and suggestions. This work is funded by the National Basic Research Program of China (Grant No. 2012CB316300), the National Natural Science Foundation of China (Grant No. 61273272, 61075024), the International S&T Cooperation Program of China (Grant No. 2010DFB14110) and the National Key Technology R&D Program (Grant No. 2012BAK02B01). Zhenan Sun is the corresponding author of this paper.

REFERENCES

Zhenhua Chai received the B.E. degree in Automation (with Honors) from Central University of Nationality (CUN) and the Ph.D. degree in Computer Application Technology from National Lab of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences (CASIA) in 2008 and 2013 respectively. His research interests focus on face recognition, feature extraction and system integration. He has published several papers on feature extraction for face recognition.

Zhenan Sun is currently an Associate Professor in Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR) at Institute of Automation, Chinese Academy of Sciences (CASIA). He received the B.E. degree in Industrial Automation from Dalian University of Technology, the M.S. degree in System Engineering from Huazhong University of Science and Technology, and the Ph.D. degree in Pattern Recognition and Intelligent Systems from CASIA in 1999, 2002 and 2006, respectively. Since March 2006, Dr. Sun has joined the NLPR of CASIA as a faculty member. Dr. Sun is a member of the IEEE, the IEEE Computer Society and the IEEE Signal Processing Society. Dr. Sun’s research areas include biometrics, pattern recognition and computer vision, and he has authored/co-authored over 100 technical papers. He is an associated editor of IEEE Transactions on Information Forensics and Security and IEEE Biometrics Compendium.

Heydi Mendez-Vazquez received the B.S. degree (with Honors) in Computing in 2005 and the Ph.D. in Automatic and Computing in 2010, all from the Polytechnic University of Havana (CUJAE), Cuba. Since 2005, she has been with the Biometric Research Group of the Advanced Technologies Applications Center (CENATAV), in Cuba. Her research interests include computer vision, biometrics, face recognition and image processing. She is a member of the Cuban Association of Pattern Recognition (ACRP) and the International Association of Pattern Recognition (IAPR). She received a best student paper award in the 5th International Summer School on Biometrics endorsed by the IAPR, in Alghero, Italy, 2008 and was awarded with one of the annual prizes given by the Cuban Academy of Sciences in 2011.

Ran He received the B.E. degree in Computer Science from Dalian University of Technology, the M.S. degree in Computer Science from Dalian University of Technology, and Ph.D. degree in Pattern Recognition and Intelligent Systems from Institute of Automation, Chinese Academy of Sciences in 2006, 2008 and 2009, respectively. Since December 2010, Dr. He has joined NLPR where he is currently Associate Professor. He currently serves as an associate editor of Neurocomputing (Elsevier) and serves on the program committee of several conferences. His research interests focus on information theoretic learning, pattern recognition, and computer vision.

Tieniu Tan is currently a Professor in Center for Research on Intelligent Perception and Computing (CRIPAC), National Laboratory of Pattern Recognition (NLPR) at Institute of Automation, Chinese Academy of Sciences (CASIA). He received his B.Sc. degree in electronic engineering from Xi’an Jiaotong University, China, in 1984, and his M.Sc. and Ph.D. degrees in electronic engineering from Imperial College of Science, Technology and Medicine, London, U.K., in 1986 and 1989, respectively. His current research interests include biometrics, image and video understanding, information hiding and information forensics. Dr. Tan is a Fellow of the IEEE and the IAPR (the International Association of Pattern Recognition).