Face Identification Using Novel Frequency-Domain Representation of Facial Asymmetry

Sinjini Mitra, Marios Savvides, *Member, IEEE*, and B. V. K. Vijaya Vijaya Kumar, *Senior Member, IEEE*

***Abstract*—Face recognition is a challenging task. This paper introduces a novel set of biometrics, defined in the frequency domain and representing a form of "facial asymmetry." A comparison with existing spatial asymmetry measures suggests that the frequency-domain representation provides an efficient approach for performing human identification in the presence of severe expressions and for expression classification. Error rates of less than 5% are observed for human identification and around 25% for expression classification on a database of 55 individuals. Feature analysis indicates that asymmetry of the different face parts helps in these two apparently conflicting classification problems. An interesting connection between asymmetry and the Fourier domain phase spectra is then established. Finally, a compact one-bit frequency-domain representation of asymmetry is introduced, and a simplistic Hamming distance classifier is shown to be more efficient than traditional classifiers from storage and the computation point of view, while producing equivalent human identification results. In addition, the application of these compact measures to verification and a statistical analysis are presented.**

***Index Terms*—Asymmetry, efficiency, expression, face, features, frequency domain, identification, one-bit code, phase.**

I. Introduction

In the modern world, the ever-growing need to ensure a system's security has spurred the growth of the newly emerging technology of biometric-based identification. Of all the biometrics that are being used today, the face is the most acceptable because it is one of the most common methods that humans use in their visual interaction and perception. In fact, facial recognition is an important human ability—an infant innately responds to face shapes at birth and can discriminate his or her mother's face from a stranger's at the tender age of 45 h [1]. In addition, the method of acquiring face images with digital cameras is nonintrusive. However, face-based identification poses many challenges. Several images of a single person may be dramatically different because of changes in viewpoint, color, and illumination, or simply because the person's face looks different from day to day due to makeup, facial hair, glasses, etc. Faces are rich in information about individual identity, mood, and mental state, and position relationships between face parts, such as eyes, nose, mouth, and chin, as well as their shapes and sizes are widely used as discriminative features for identification. One family of features that has only recently come into use in face recognition problems is facial asymmetry.

Facial asymmetry can be caused either by external factors such as expression changes, viewing orientation and lighting direction, or by internal factors such as growth, injury, and age-related changes. The latter is more interesting, being directly related to the individual face structure, whereas the former can be controlled to a large extent and even removed with the help of suitable image normalization. Psychologists have long been interested in the relationship between facial asymmetry and attractiveness and its role in identification. It has been observed that the more asymmetric a face is, the less attractive it is [2]. Furthermore, the less attractive a face is, the more recognizable it is [3]. A commonly accepted notion in computer vision is that human faces are bilaterally symmetric [4]. Gutta *et al.* [5] reported no difference whatsoever in recognition rates while using only the right and left halves of the face. However, a well-known fact is that manifesting expressions cause a considerable amount of facial asymmetry, as they are more intense on the left side of the face [6]. Differences were indeed found in recognition rates for the two halves of the face under a given facial expression [7]. All of these indicate the potential significance of asymmetry in automatic human face recognition, particularly in the presence of expressions. Identifying twins based on face images has defied the best of facial recognition systems, and [8] reported statistically significant differences among facial asymmetry parameters of monozygotic twins. This shows the potential of facial asymmetry in producing efficient identification tools.

Despite extensive studies on facial asymmetry, its use in human identification began in computer vision in 2001 with the seminal works by Liu [9], who, for the first time, showed that certain facial asymmetry measures are efficient human identification tools under expression variations. This was followed by more in-depth studies [10], [11] which further investigated the role of asymmetry measures both for human and expression classifications. The goal in this paper is to study an alternative representation of facial asymmetry in the frequency domain, which constitutes a completely novel and unique research agenda. We wish to establish its efficacy in various recognition tasks and other properties by exploiting the characteristics of the frequency domain.

The rest of the paper is organized as follows. Section II contains a brief description of the database and Section III introduces the new frequency-domain asymmetry biometrics. Section IV presents a feature analysis, and the classification results appear in Section V. Section VI explores the connection

Manuscript received October 24, 2005; revised May 19, 2006. This work was supported by a grant from the Army Research Office to CyLab, Carnegie Mellon University. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Anil Jain.

S. Mitra is with the Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292 USA (e-mail: mitra@isi.edu).

M. Savvides and B. V. K. Vijaya Kumar are with the Electrical and Computer Engineering Department and CyLab, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: msavvid@cs.cmu.edu; kumar@ece.cmu.edu).

Digital Object Identifier 10.1109/TIFS.2006.879301

Neutral
Joy
Disgust
Anger

Fig. 1. Sample images from our database. (Courtesy Liu et al. [13]).

between asymmetry and Fourier domain phase spectra while Section VII introduces a computationally efficient one-bit code. We conclude with a discussion in Section VIII.

II. Data

We use a part of the "Cohn–Kanade AU-coded Facial Expression Database" [12], consisting of images of 55 individuals expressing three different emotions—joy, anger, and disgust. The data thus consist of video clips of people showing an emotion, beginning with a neutral expression and gradually evolving into its peak form. Each clip is broken down into several frames and the raw images are normalized using an affine transformation technique based on a combination of scaling and rotation that was employed by [13]. Each normalized image is of size 128×128 and has a face midline determined so that every point on one side of the face has a corresponding point on the other. We do not include details here for space constraints but an interested reader is referred to [13] for details on the alignment procedure. Some normalized images from our dataset are shown in Fig. 1. This is the only known database as per our knowledge, that allows for a thorough investigation of the role of facial asymmetry in identification in the presence of extreme expression variations, since the images were carefully captured under controlled background lighting. We use this small subset as the initial testbed for our experiments and hope to extend to a bigger database in the near future. This also facilitates a fair comparison of our results to those in [9]–[11], which were based on this small subset too.

III. Frequency Domain

Many signal processing applications in computer engineering involve the frequency-domain representation of signals. The frequency spectrum consists of two components at each frequency: magnitude and phase. In two-dimensional (2-D) images particularly, the phase component captures more of the image intelligibility than magnitude and, hence, is very significant for performing image reconstruction [14]. Savvides *et al.* [15] showed that correlation filters built in the frequency domain can be used for efficient face-based recognition. Recently, the significance of phase has also been used in biometric authentication. Savvides *et al.* [16] proposed correlation filters based only on the phase component of an image, which performed as well as the original filters and [17] demonstrated that performing principal component analysis (PCA) in the frequency domain by eliminating the magnitude spectrum and retaining only the phase not only outperformed spatial-domain PCA, but also have attractive properties such as illumination tolerance. All of these show that frequency-domain features possess the potential for improving classification results.

Symmetry properties of the Fourier transform are often very useful. According to [18], any sequence $x(n)$ can be expressed as a sum of an even part or the symmetry part $x_e(n)$ and an odd part or the asymmetry part $x_o(n)$. Specifically

$$x(n) = x_e(n) + x_o(n)$$

where $x_e(n) = 1/2(x(n) + x(-n))$ and $x_o(n) = 1/2(x(n) - x(-n))$. When a Fourier transform is performed on a real sequence $x(n)$, the even part $(x_e(n))$ transforms to the real part of the Fourier transform and the odd part $(x_o(n))$ transforms to its imaginary part (Fourier transform of any sequence is generally complex-valued). In the context of a face image, the even part corresponds to the symmetry of a face (in the left–right direction, across from the face midline) and, hence, the more symmetric a face region is, the higher the value will be of the corresponding odd part and vice-versa. This implies that spatial asymmetry of the face corresponds to the imaginary part of the Fourier transform and the symmetry part corresponds to the real part, and this correspondence lays the ground for developing asymmetry features in the frequency domain. However, all of these relations hold for one-dimensional (1-D) sequences alone and, hence, we define our asymmetry features based on the Fourier transforms of row slices of the images (either singly, or averaging over a certain number of rows at a time, as described in the next section).

A. Asymmetry Biometrics

Following the notion presented above, we define three asymmetry biometrics in the frequency domain based on the imaginary components of the Fourier transform as:

- *I-face:* frequency-wise imaginary components of Fourier transforms of each row slice—128×128 matrix of features (need to use half of these owing to symmetry properties of Fourier Transform—128×64 features);
- *Ave I-face:* frequency-wise imaginary components of Fourier transforms of averages of two-row slices of the face—64×64 matrix of features;
- *E-face:* energy of the imaginary components of the Fourier transform of averages of two-row slices of the face—a feature vector of length 64.

For all three sets of features, the higher their values, the greater the amount of asymmetry, and vice-versa. The averaging over rows is done in order to smooth out noise in the image which is likely to create artificial asymmetry artifacts and give misleading results. Averaging over more rows, on the other hand, can lead to oversmoothing and a loss of relevant information. The two-row blocks were selected as optimal after some experimentation. We also wish to compare the performances of the three feature sets, especially to explore whether it is justified to consider the higher-dimensional I-faces instead of the E-faces in terms of better classification performance. Note that all of these features are simple in essence, yet the goal is to show that they are capable of forming effective identification tools. To the best

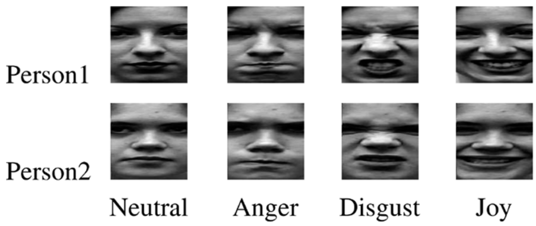

Fig. 2. Two people showing four different facial expressions.

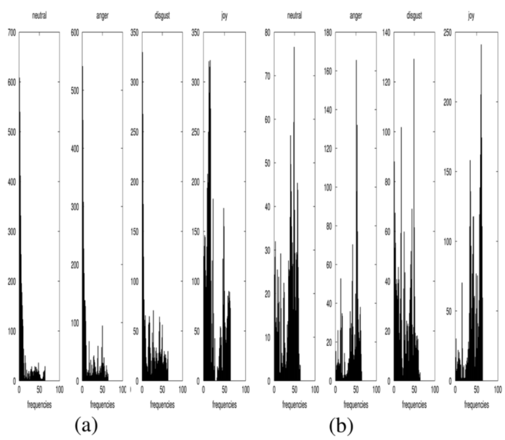

Fig. 3. Asymmetry of the different facial features for the four expressions of two people. The horizontal axis represents the different frequencies at which the asymmetry biometrics were computed for each row slice of the face from the forehead to the chin. (a) Person 1. (b) Person 2.

of our knowledge, these frequency-based features representing facial asymmetry are fairly novel in any computer vision and pattern recognition problem.

IV. Feature Analysis

The different exploratory feature analyses that we present in this section are aimed at providing a preliminary idea about the nature of the frequency asymmetry metrics and their utility in classification methods. For the first set of feature analysis, we use the E-faces due to their low dimensionality. Fig. 2 shows the images of two individuals with four different expressions, and Fig. 3 shows how asymmetry varies among them. For instance, for person 1, joy produces the greatest degree of asymmetry, and neutral expression the lowest, whereas for person 2, joy and neutral expressions show maximum asymmetry followed by anger and disgust. Moreover, Person 1 has a greater amount of asymmetry over the whole face for joy (forehead, nose, and mouth regions) while only in the forehead region for the other three emotions. As for Person 2, on the other hand, the mouth region appears to have the maximum asymmetry for all four emotions. Therefore, although we looked at only two people in the database, these analyses give a preliminary idea that people may tend to express different emotions differently which, in turn, suggests that these measures may be helpful in automatic face recognition tasks in the presence of expression variations as well as in identifying expressions.

We next study the distribution of the asymmetry metric for all of the 55 people for certain facial regions that not only contain

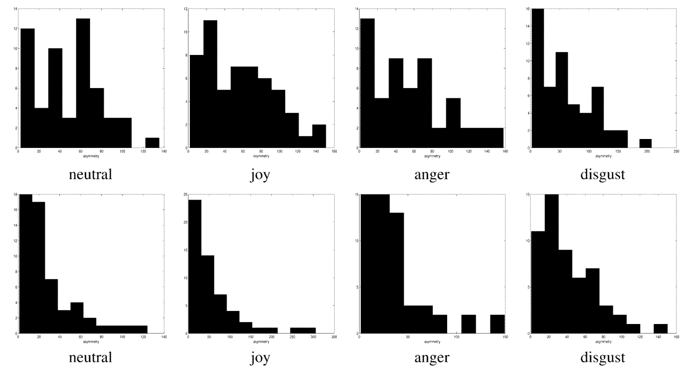

Fig. 4. Distributions of the asymmetry metric for 55 people and for the four different expressions. The top panel shows the distributions of the eye region and the bottom panel corresponds to the mouth region.

valuable discriminating information but also are prone to significant changes during expressing emotions—eyes and mouth. We select a few rows in those regions, average them, and compute the energy of the imaginary parts of the Fourier transform (measure of asymmetry). Fig. 4 shows these energy values for the four expressions for all 55 people in our dataset. The multiple peaks in these figures reveal that there exists a significant difference in asymmetry among the different people and also among the four expressions.

A. Discriminative Feature Sets

Next, we study the discriminative power of the asymmetry measures to determine the specific facial regions that contribute to the identification process, both for humans and expressions. Ideally, those features which contribute to interclass differences should have a large variation between classes and small variation within the same class. Hence, a measure of discrimination can be provided by a variance ratio type quantity; in particular, we use what is known as an augmented variance ratio (AVR), following along the lines of [13]. AVR compares within class and between class variances and, at the same time, penalizes features whose class means are too close to one another. For a feature F with values S_F in a data set with C total classes, AVR is calculated as

$$\mathrm{AVR}(S_F) = \frac{\mathrm{Var}(S_F)}{\frac{1}{C}\sum_{k=1}^{C} \frac{\mathrm{Var}_k(S_F)}{\min_{j\neq k}(|\mathrm{mean}_k(S_F) - \mathrm{mean}_j(S_F)|)}}$$

where $\mathrm{mean}_i(S_F)$ is the mean of the subset of values from feature F belonging to class i. The higher the AVR value of a feature is, the more discriminative it is for classification. For human identification, the 55 subjects form the classes and for expression classification, the classes are the three emotions.

Fig. 5 shows the E-face AVR values for human and expression classifications calculated based on all 55 individuals in the database. Looking carefully at the human AVR values, we discover that a few subjects in the database have some artificial asymmetry in the forehead region arising from either falling hair or edge artifacts introduced in the normalization procedure (Fig. 6). This is highly undesirable and spuriously raises the first few AVR values in Fig. 5(a). Fig. 5(b) shows the AVR plot

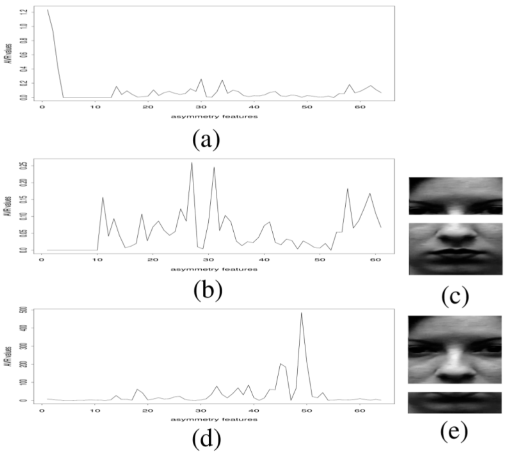

Fig. 5. AVR values for E-faces. Human identification: (a) for all features, (b) all features except top 3, (c) nose bridge with the highest AVR value. Expression classification: (d) all features and (e) the mouth region with the highest AVR value. The features 0–64 represent the regions from the forehead to the chin of a face.

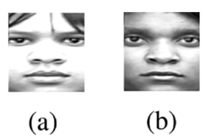

Fig. 6. Images with artificial asymmetry in the forehead. (a) Hair. (b) Edge artifacts.

without the top three features, and this clearly shows that the nose bridge contains the most discriminative information pertaining to recognition of individuals under different expressions [marked in Fig. 5(c)]. Fig. 5(d)–(e) shows that the region around a person's mouth is most discriminative across different expressions (no problem here due to edges). We thus conclude that the asymmetry of different face regions drives these two apparently conflicting classification problems and, hence, may be effective for both. Moreover, these results are consistent with similar feature analysis results in [11], based on spatial asymmetry measures.

V. Classification Results

We tried different classification methods which include Fisher faces (FF) [19], support vector machines (SVMs) [20], linear discriminant analysis (LDA) [21], and the individual principal component analysis (IPCA) [22]. The IPCA method is different from the global PCA approach [23] where a subspace W is computed from all of the images regardless of identity. In an individual PCA, on the other hand, subspaces W_p are computed for each person p and each test image is projected onto each individual subspace using $y_p = W_p^T(x - m_p)$. The image is then reconstructed as $x_p = W_p y_p + m_p$ and the reconstruction error computed $||e_p||^2 = ||x - x_p||^2$. The final classification chooses the subspace with the smallest $||e_p||^2$. Of these four classifiers, LDA did not perform well and so we omit those results and report those from the other three classifiers. LDA is known to be effective for spatial or image domain features computed from pixel intensity values, and this may be a probable cause of its failure for our frequency-based features.

TABLE I
Misclassification Rates for Human Identification Using Frequency-Domain Asymmetry Measures

Classifiers	I-face	Ave I-face	E-face
FF	15.76%	10.91%	19.19%
SVM (Linear kernel)	4.85%	3.94%	8.79%
SVM (polynomial, deg 2)	6.36%	6.36%	9.70%
IPCA	4.85%	3.64%	6.36%

TABLE II
Misclassification Rates for Expression-Invariant Human Identification Based On Spatial D-Face Measures

Asymmetry features	Error rates	Asymmetry features	Error rates
Spatial D-face	17.58%	Spatial D-face PCs	3.03%

A. Human Identification

For human identification, we train on the neutral frames of the three emotions of joy, anger, and disgust from all 55 individuals, and test on the peak frames of the three emotions from all of the people. We thus use three frames per person for training (165 total) and three frames per person for testing (165 total). This represents an expression-invariant human identification problem, similar to the one reported in [13] which uses a spatial asymmetry measure called Difference face (D-face) defined as $D(x,y) = I(x,y) - I'(x,y)$. I denotes a (normalized) face image and I' is its reflected version along the face midline. LDA was used as the classifier in their case. Note that [13] reported classification results on five different experiments—training on frames from two emotions and testing on the third, training on neutral frames and testing on peak ones and vice-versa. However, we use only one of these experimental setups (training on neutral and testing on peak) and, hence, we will compare our results to the corresponding cases from the spatial measures.

Table I shows the misclassification error rates for human identification (percentage of cases that are wrongly classified) based on our frequency-domain asymmetry biometrics from the different classifiers. They show that IPCA performed the best, closely followed by linear-kernel-based SVM (in fact, their I-face and Ave I-face results are almost identical to the IPCA ones). Moreover, I-faces proved to be significantly better than E-faces, and this shows that the feature reduction by way of summing over each row in constructing the E-faces destroyed features crucial for discrimination and, hence, deteriorated performance. We will henceforth work with the I-faces alone.

We next compared our best results from IPCA with those obtained with D-faces [13] shown in Table II. We compare to D-face alone (except for D_{hx} which are significantly worse) because, by construction, I-faces and E-face are analogous to those (following the reasoning presented in Section III). The results indicate that our proposed frequency-domain measures are significantly better than D-face and have no statistically significant differences with the D-face principal components (PCs) at the 1% level. These are adjudged by using "p-values" (or probability values), a quantity commonly used in statistical inference problems to test the significance of a proposed hypothesis. The lower the p-value of a test, the more evidence the data exhibit against its acceptance. (Casella [24] contains details on the procedure of statistical hypothesis testing). For both I-faces and

TABLE III
MISCLASSIFICATION RATES FOR EXPRESSION CLASSIFICATION. THE FIGURES IN THE PARENTHESES DENOTE THE STANDARD DEVIATIONS OVER THE 20 REPETITIONS

Features	Misclassification rates
I-face (IPCA)	26.93% (4.18%)
Ave I-Face (IPCA)	27.07% (3.77%)
I-face (SVM, linear)	38.90% (2.53%)
Ave I-face (SVM, linear)	37.45% (5.07%)
D-face	39.60% (2.74%)
D (PCs)	36.73% (3.88%)

TABLE IV
MISCLASSIFICATION RATES FOR HUMAN IDENTIFICATION AND EXPRESSION CLASSIFICATION USING R-FACES. STANDARD DEVIATIONS FOR THE LATTER APPEAR IN PARENTHESES

Classification	Features	Error rates
Human	S-face	29.70%
	S-face (PCs)	10.91%
	R-face	10.30%
Expression	S-face	18.67% (3.22%)
	S-face (PCs)	17.80% (3.88)
	R-face	27.07% (6.31%)

TABLE V
ERROR RATES FOR HUMAN IDENTIFICATION AND EXPRESSION CLASSIFICATION USING I-FACE+R-FACE COMBINATION FEATURES. STANDARD DEVIATIONS FOR THE EXPRESSION CLASSIFICATIONS ARE COMPUTED OVER 20 REPETITIONS

Classification	Feature sets	Error rates	Std. dev.
Human	Ave I-face+R-face	2.78%	-
Expression	I-face+R-face	18.90%	3.69%
Human	D-face+S-face	9.70%	-
	D-face (PC) + S-face (PC)	1.21%	-
Expression	D-face+S-face	16.60%	5.92%

E-faces, we test the hypothesis that the error rates from using the spatial-domain and the frequency-domain measures are the same. When comparing D-face, we obtain p-values < 0.0001 and with D-face PCs, p-values > 0.01, thus suggesting that the frequency-domain measures are at least as robust to intrapersonal variations caused by expression changes as their spatial-domain counterparts.

B. *Expression Classification*

A person's expression is helpful to identify his or her mood and mental state, and is often an individualized characteristic. Different people express different emotions differently, which echoes human behavior and often helps in the identification of a particular individual. In fact, [7] showed convincing results that face-recognition rates depend on different types of facial expressions.

Our dataset has images with three different expressions: joy, anger, and disgust. We follow the same experimental setup as in [11]—train on peak frames from all three expressions for a randomly selected subset of 30 individuals (out of a total of 55), and test on peak frames of the three expressions from the remaining 25 individuals. This random division of the subjects into training and test sets was repeated 20 times (in order to remove selection bias) and the final error rates are obtained by averaging over those from these 20 repetitions.

The results tabulated in Table III show that unlike the case of human identification, the frequency-domain features have lower misclassification rates than both the D-face measures (again, comparing only with D-face) with significant improvements of over 10% (p-values< 0.006 in all cases) with IPCA. SVM, however, proved to be not as efficient for expression classification as it was for human identification, although the results were not significantly worse than those based on D-face (but significantly worse than IPCA). The FF results are considerably poorer and, thus, have been ignored. Note that all of these results, although not very satisfactory, are at least significantly better than those obtained from pure random guessing (probability of error: 66.67% for a three-class problem).

C. *"Edge"-Based Features*

Edges are known to contain information valuable for discriminating among individuals and, hence, it seems natural to consider identification features based on those. Mitra [11] and Liu *et al.* [13] used a set of spatial asymmetry measures called symmetry face (S-face) defined as $S(x,y) = \cos(\phi_{I_e(x,y),I'_e(x,y)})$, where I_e is the edged image of I, and I'_e is its vertically reflected image. Just as S-faces were designed to emphasize symmetry instead of asymmetry (higher values indicate more symmetry, see [13]), we construct a set of frequency-domain symmetry measures using the real parts of the Fourier transforms of the one-dimensional (1-D) row slices of the edged images I_e. Recall here that the symmetric part of any sequence transforms to the real part of Fourier transform (Section III). We call these R-faces and the higher its value for a feature, the more symmetric (less asymmetric) the corresponding facial region is and vice-versa.

The same experimental setup for human identification and expression classification as with the I-faces are followed here. The results in Table IV (only the best ones with IPCA) show that the R-face yields lower error rates than the spatial S-face for human identification (we only compare the spatial S-face results owing to the correspondence, and omit the S_{hx} results which were quite poor). These are, however, higher than I-face ones (Table I), and this is consistent with what [13] observed (D-face results better than S-face results). For expression classification, on the other hand, the error rates are poorer than those obtained with S-face, but they are almost identical to those obtained with I-face. However, the spatial S-face features were more efficient than D-face features for expression classification [11]. We thus conclude that R-faces are useful for identifying people in the presence of expressions, but not so much for classifying expressions when compared with the corresponding spatial measures.

D. *Combining R-Faces With I-Faces*

In order to investigate whether the I-face and the R-face feature sets complement each other for better performance, we concatenate the two sets of features to yield a 2-D feature vector per frequency and perform both human identification and expression classification using the same setup as before. We use the Ave I-faces for human identification and the I-faces for expression classification, since they give the best results for the respective problems, and IPCA classifier. The results from the combination of the feature sets in Table V indicate that improvement occurs in their performances for both classification tasks.

TABLE VI
RELATIVE IMPROVEMENTS IN THE ERROR RATES OVER THE INDIVIDUAL FEATURE SETS BY COMBINING THEM

	I-face	R-face	I-face+R-face	Imp.(I-face)	Imp.(R-face)
Human	3.64%	10.30%	2.78%	23.63%	73.01%
Expr.	26.93%	27.07%	18.90%	29.82%	30.18%

For human identifiation, the R-face results improve significantly whereas the amount of improvememnt in the I-face results is limited which is reasonable given the already good results from using them alone. As to expression classification, both the feature sets show significant improvements as a result of the combination. On comparing with the combination of the S-faces and D-faces (spatial domain), we find that the combination of I-faces and R-faces does not perform as well as the combination of the D-face and the S-face PCs. However, it outperforms the D-face PCs alone for human identification and they are very nearly equal to the S-face PCs (alone) for expression classification. Table VI shows the improvements achieved by the combination over the individual feature sets relative to their original results as was done for the spatial feature sets, calculated as

$$\text{Imp.} = \frac{\text{Individual error rate - Combined error rate}}{\text{Individual error rate}} \times 100.$$

VI. CONNECTION WITH PHASE

As mentioned in Section III, the Fourier domain phase is important for face identification, and this section investigates a potential connection between phase and facial asymmetry. Recall that the symmetric part of a sequence transforms to the real part in a Fourier transform whereas the asymmetric part transforms to the imaginary part. Thus, the Fourier transform of a real symmetric sequence is real; that of a real asymmetric sequence is purely imaginary. The phase angle is defined as

$$\theta = \tan^{-1}\left(\frac{I}{R}\right)$$

where R and I are, respectively, the real and the imaginary parts of the Fourier transform. So $\theta = 0$ if and only if the imaginary component I is zero. In other words, a completely symmetric sequence gives rise to zero-phase frequency spectrum. In other words, the absence of asymmetry implies zero phase and vice-versa. Note here that that all of the above relations hold for 1-D sequences only (that is, when performing 1-D Fourier transforms), as also pointed out in Section III.

Interpreting these relations in terms of a face image, a completely symmetric face does not necessarily imply a zero-phase spectrum when applying a 2-D Fourier transform. However, if we treat each row slice of the face as a 1-D sequence and apply a 1-D Fourier transform, the above relations hold and each symmetric row will give rise to a 1-D zero-phase spectrum. To make this study more rigorous, we constructed similar asymmetry features as before, using 1-D row slices or averages of two-row slices but now based on "phase-only images" which have constant unit magnitude (obtained by dividing the Fourier transform of an image by its magnitude). We call them I-face$_\theta$ and Ave I-face$_\theta$ to distinguish them from the I-faces from before. Fig. 7 shows a phase-only image corresponding to an original face image. A close look at the figure shows that a phase-only image retains most of the relevant facial features (eyes, nose, mouth) and, hence, much of the image identifiability.

(a) (b)

Fig. 7. (a) Original Image. (b) Phase-only image.

TABLE VII
MISCLASSIFICATION RATES FOR HUMAN AND EXPRESSION CLASSIFICATIONS USING PHASE-ONLY IMAGES. STANDARD DEVIATIONS FOR THE LATTER OVER THE 20 REPETITIONS APPEAR IN PARENTHESES

Features	Human	Expr.
I-face$_\theta$	4.85%	37.8% (4.98%)
Ave I-face$_\theta$	5.45%	38.8% (4.76%)

Human and expression identification results using the same experimental setup as before (training on neutral frames and testing on peak frames using the IPCA classifier), but with the phase-only images, are shown in Table VII. The human identification results help establish, although empirically, an interesting connection between facial asymmetry and phase for the purpose of human identification. The I-face features based on the actual images and the phase-only images produced exactly the same human classification results. This indicates that phase contains all of the asymmetry of the original face, at least to the extent that is necessary for human classification purposes, and no crucial information is lost by removing the magnitude. In the other direction, we conclude that the classification based on asymmetry might have been driven by the phase information contained in it. This leads us to believe that asymmetry may provide an alternative means of representing phase information. On the other hand, this explains the success of asymmetry measures in human identification in terms of its direct relationship with phase, since the vital role of the latter in human recognition is well known. The Ave I-face$_\theta$ results are a little worse than the Ave I-face results (not significantly though) and this may have resulted from the averaging involved in their construction. However, the expression classification results are considerably worse than before, which shows that removing the magnitude destroys relevant information for identifying expressions.

VII. IMPROVING EFFICIENCY: ONE-BIT CODE

We have seen that frequency-based asymmetry measures are very capable identification tools. We now focus on improving upon the efficiency of these measures, in terms of storage and computational requirements. This is a primary consideration for biometric systems in practice. For the time being, we do this in the context of human identification alone, but can be similarly

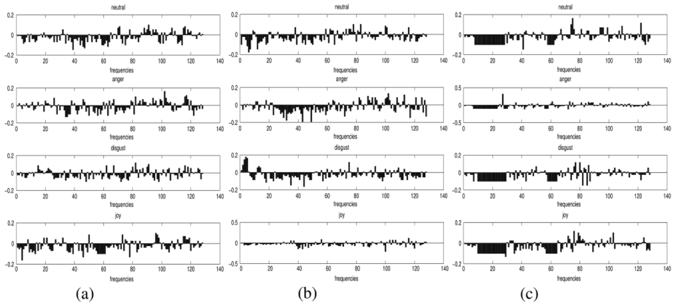

Fig. 8. FAC for the four expressions of three individuals. For the top four rows, +ve values denote more asymmetry and -ve values more symmetry. The x-axis represents the parts of the face from the forehead to the chin while the y-axis shows the rowwise averaged FAC values. (a) Person 1. (b) Person 2. (c) Person 3.

applied to expression classification as well. To this end, we define a new set of features as follows, for each frequency x of a 1-D Fourier transform

$$F(x) = \begin{cases} +1, & \text{if } I(x) > R(x) \\ -1, & \text{if } I(x) \leq R(x) \end{cases}$$

where $I(x)$ and $R(x)$, respectively, denote the imaginary and the real part of the Fourier frequency x. Since each such feature is of one bit per frequency, we call them one-bit facial asymmetry code (FAC). What the features describe is as follows: for a particular frequency, $F(x) = 1$ implies that the frequency represents more asymmetry than symmetry, and vice-versa if $F(x) = -1$. It is a very compact representation of asymmetry and the features are easy to compute and store, requiring much less memory than usual quantified measures.

We consider two sets of features: 1) the frequency-wise FAC values—128×128 matrix and 2) FAC computed on Fourier transforms of two-row averages of the original image—64×128 matrix, denoted ave FAC.

A. Feature Analysis

For the exploratory feature analysis (and only for this), we consider a reduced dimension FAC set constructed as follows: the frequency-wise FAC bits are averaged over each row, so that if $b(x, y)$ denotes the bit at frequency (x, y), we compute $B(x) = 1/N \sum_y b(x, y)$ where N denotes the number of frequencies in each row (or the number of columns). This means that if $B(x) > 0$ for a particular row, the features in that row are more asymmetric and if $B(x) < 0$, the features in that row are more symmetric. This feature reduction technique helps in reducing the noise of the frequency-wise values and facilitates studying the feature patterns more clearly. Fig. 8 shows the pattern of the variation of FAC for three people while expressing different emotions (one frame per emotion). They give a preliminary but convincing idea that these measures may be helpful in recognizing people in the presence of expression variations owing to the existence of somewhat distinct patterns for each person.

As with the E-faces earlier, we also computed the AVR values of the FAC features and observed that the region above the eyes contains the most discriminative information followed by the nose bridge region. We do not include the plot here due to space constraints.

TABLE VIII
MISCLASSIFICATION ERROR RATES FOR THE FAC-BASED FEATURES

Classifiers	FAC	Ave FAC
IPCA	4.24%	4.54%
SVM (linear kernel)	3.94%	5.15%
SVM (polynomial, deg 2)	3.64%	5.15%
HD	4.24%	4.54%

TABLE IX
STORAGE REQUIREMENTS OF HD, PCA (EIGENVECTORS) AND SVM (SUPPORT VECTORS) CLASSIFIERS FOR IMAGES OF DIFFERENT SIZES

Actual size	Image storage (bits)	PCA/SVM (bits)	HD (bits)
64×64	32768	131072	2048
128×128	131072	524288	8192

B. Human Identification Results

As before, we again train on the neutral frames from all of the people and test on the peak frames of the three emotions. Since our features are essentially encoded as bit patterns, it seems natural to use a distance-type metric that is more effective for comparing bit patterns. Once such metric is the popular Hamming distance (HD), which gives the count of bits that are different in two patterns. More generally, if two ordered list of items are compared, HD is the number of items that do not match identically. In our case, when comparing two FAC patterns, HD outputs the number of bits in two codes that do not match. We use this classifier in addition to IPCA and SVM that we used earlier.

The misclassification error rates appear in Table VIII which show that our proposed FAC results are as good as the I-face results reported in Section V and there is no statistically significant difference between them. The lowest error rates are obtained with SVM, although the HD results are not significantly different from them. Since FAC is much more compact than the I-face representation, this implies a huge gain in terms of efficiency. Just as the I-faces, these results are significantly better than the original D-face results and no worse than the D-face PC results (Table II). Besides, the HD classifier also has a definite advantage over PCA and SVM-based methods with respect to computational resources. Its computation is much less intensive (involves Boolean exclusive-OR operation only) and is much simpler to store than the eigenvectors of PCA or the support vectors of SVM, which require floating point 32-b representation. Moreover, only half of these codes need to be stored and used due to the conjugate Hermitian symmetry arising from purely real sequences, according to which frequencies are symmetric around the origin (the real part is symmetric and imaginary part is odd-symmetric). So, in essence, we are just using half-bit codes in the matching routine. A comparison of the storage requirements shown in Table IX shows that HD requires up to 64 times less storage space than either PCA or SVM for operation and even 16 times less storage space than the normalized images themselves. This alone establishes a firm basis for the utility of the HD classification algorithm based on FAC for performing face recognition in practice. Efficiency is definitely a crucial factor in classification problems in computer vision since it usually involves high-dimensional images. So, based on the results in this section, we conclude that not only are the FAC features more efficient than the I-faces in terms of representation, but also the HD classifier is more capable than traditional PCA and SVM from storage and computational standpoints, without any compromise on the classification performance.

C. Extension to Verification

The HD-based classifer presents the scope for carrying out statistical analysis of the underlying matching mechanism which shows that it may be useful for verification tasks. If X is a random variable denoting the number of matched bits for a pair of FACs using HD, then assuming that the individual bits are uncorrelated, X follows a binomial distribution with parameters p (probability of a match) and n (total number of bits per image). Now, if Y_i denotes the total number of matched bits for person i when matching N_i images of this person, then $Y_i = \sum_{k=1}^{N_i} X_k^i, X_k^i$ is the number of matched bits for the k^{th} image of person i. Then, $Y_i \sim \text{Bin}(nN_i, p_i), i = 1, \ldots, 55$, where p_i is the probability of a matched bit for person i. $p_i = p$ implies that every person has the same probability of match per bit. Note that HD gives the number of mismatched bits for a pair of FAC, say Z, then $X = n - Z$.

We estimate p_i by the sample proportions of match, given by $\hat{p}_i = y_i/nN_i, i = 1, \ldots, 55$. The 95% confidence interval for each p_i is then given by $\hat{p}_i \pm 1.96 \times \hat{\sigma}_i$, where $\hat{\sigma}_i = \sqrt{\hat{p}_i(1-\hat{p}_i)/nN_i}$ using the normal approximation to binomial since we have a large number of samples. We compute these estimates for two cases.

Case 1) "Authentic," when matching two FACs from the same person.

Case 2) "Impostor," where two FACs belonging to two different people are matched.

The distribution of Y_i in Fig. 9 shows that the number of matched bits is much higher for the authentic cases (> 1500 for most cases) than the impostor ones (less than 1000) which is what one expects. Fig. 10 shows the sample estimates of the matching probabilities of the authentics and the impostors for the 55 individuals in the database, along with the associated 95% confidence intervals. As expected, the estimates for the genuine cases are considerably higher than the impostor ones. However, the upper confidence limits for the latter seem a little higher than desirable (greater than 0.5 in some cases). This is attributed to inflated standard deviations caused by variation in the impostor probabilities across people. This happens because some people are more identical looking and, hence, more likely to be mistaken for each other than others. One way to rectify this will be to consider impostor probabilities for pairs of people taken at a time. Statistical tests also reveal the existence of significant differences between the bit-matching probabilities of the authentics and impostors for all of the people.

Motivated by the above statisitical analysis, we now use the HD classifier and FAC-based asymmetry features for verification purposes. Verification is performed using a threshold on the number of matched bits per test image and Fig. 11 shows the corresponding ROC curve which plots the false acceptance rate (FAR) versus the true acceptance rate. We observe an equal error rate (EER) as low as 2% at an optimal threshold of 1500 matched bits (uniformly for all people) where the FAR equals the false rejection rate (FRR). Thus, the FAC features produce better verification results than classification results, which occurs since verification is an easier task involving only 1-to-N comparisons as opposed to N-to-N comparisons for classification.

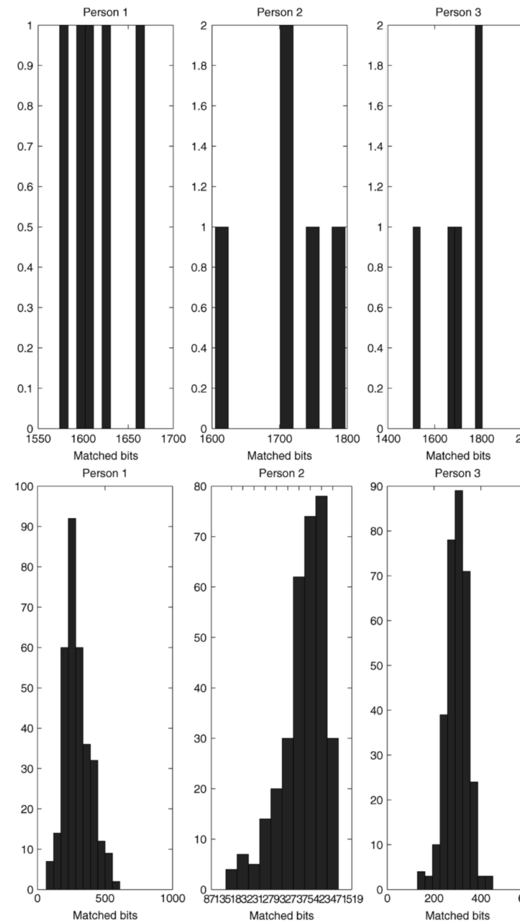

Fig. 9. Number of matched bits for three people: authentic (top) and impostor (bottom).

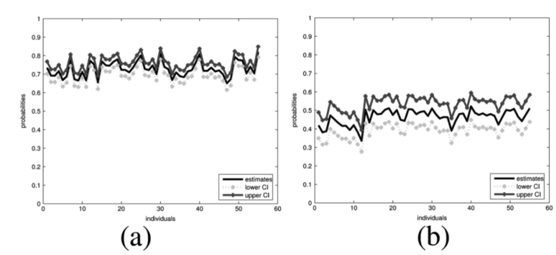

Fig. 10. $\hat{p}_i$ and the 95% confidence intervals for all 55 people. (a) Genuine. (b) Impostor.

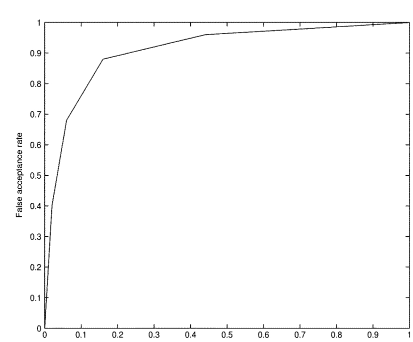

Fig. 11. ROC curve for authenticating the images in the Cohn–Kanade database, using the FAC-based asymmetry features.

VIII. Discussion

We have shown in this paper that facial asymmetry measures in the frequency domain offer a promising potential as a useful biometric in practice, especially in the presence of expression variations in face images. An error rate of less than 5% for human recognition under expression changes is impressive and desirable given that the test images are very different from the training ones. This, in turn, is very important for recognition routines in practice, for example, in biometric authentication applications where surveillance photos captured at airports are expected to be quite diverse with respect to facial expressions. Hence, any algorithm that can deal with such variations is supposed to be attractive to users from a practical point of view.

Moreover, our features are very simple and easy to compute, and they are based on well-known techniques such as Fourier transforms, phase, and edges, which are efficient face-recognition tools in computer vision. It helped show that facial asymmetry, which has so far been treated only in the spatial or image domain, also has an analogous frequency-domain representation—a fact that can be utilized in signal processing applications too. We have also established an interesting connection between facial asymmetry and the Fourier domain phase. Given the significance of phase in face-based identification, this helps in strengthening the scientific basis for the success of facial asymmetry in distinguishing human beings. This relationship is a novel one and has not been explored prior to this. It adds a whole new dimension to the concept of facial asymmetry and lays the ground for much further research in different directions. However, despite this relationship, we need to keep in mind that the frequency information is required as that explicitly describes phase information that can be used for discrimination. Such a connection could not have been established with the help of spatial-domain representation of asymmetry.

Our proposed one-bit FACs and the HD classifier also proved to be very efficient in regard to computation and storage requirements. In fact, much more than other features (D-face, I-face) and other classifiers (PCA, SVM) demonstrated. In fact, FAC needs less space than the actual gray-scale intensity images. Also, instead of storing or transmitting those, one can compute their FAC and transmit them. This is very useful for mobile, low-bandwidth communication channels, and low-memory devices, such as smart-cards and system-on-chip implementations. They yielded an error rate of less than 5% which also increases their utility even more.

One potential constraint on asymmetry features is that they require precise preprocessing of the database images. However, some preprocesing steps are necessary in most face-recognition routines since real images are usually noisy. In addition, asymmetry requires a normalization procedure to determine the face midline which acts as the reference point for computing the feaures. However, this is true for any asymmetry metric to ensure that we measure the actual intrinsic asymmetry of the face and not the artificial asymmetry artifacts created by improper normalization. In fact, we use the same detection and normalization steps as done for the spatial asymmetry measures in earlier work. It is therefore likely that our method of defining the asymmetry biometrics may be sensitive to the normalization procedure and the detection of the face midline. Hence, caution should be exercised in determining this line as accurately as possible which should not be a difficult task.

Some future research directions include the extension to a bigger database and to video streams.

References

[1] D. Voth, "In the news: Face recognition technology," *IEEE Intell. Syst.*, vol. 18, no. 3, pp. 4–7, May/Jun. 2003.

[2] R. Thornhill and S. W. Gangstad, "Facial attractiveness," *Trans. Cognit. Sci.*, vol. 3, no. 12, pp. 452–460, 1999.

[3] N. F. Troje and H. H. Buelthoff, "How is bilateral symmetry of human faces used for recognition of novel views?," *Vision Res.*, vol. 38, no. 1, pp. 79–89, 1998.

[4] S. M. Seitz and C. R. Dyer, "View morphing," in *SIGGRAPH*, 1996, pp. 21–30.

[5] S. Gutta, V. Philomin, and M. Trajkovic, "An investigation into the use of partial-faces for face recognition," in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, 2002, pp. 33–38.

[6] J. D. Borod, E. Koff, S. Yecker, C. Santschi, and J. M. Schmidt, "Facial asymmetry during emotional expression: Gender, valence and measurement technique," *Psychophys.*, vol. 36, no. 11, pp. 1209–1215, 1998.

[7] A. M. Martinez, "Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 6, pp. 748–763, Jun. 2002.

[8] P. H. Burke and M. J. Healy, "A serial study of normal facial asymmetry in monozygotic twins," *Ann. Human Biol.*, vol. 20, no. 6, pp. 527–534, 1993.

[9] Y. Liu, R. L. Weaver, K. Schmidt, N. Serban, and J. Cohn, Facial Asymmetry: A New Biometric Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-RI-TR-01-23, 2001.

[10] Y. Liu and S. Mitra, Experiments With Quantified Facial Asymmetry for Human Identification Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-RI-TR-02-24, 2002.

[11] S. Mitra and Y. Liu, "Local facial asymmetry for expression classification," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Washington, DC, Jul. 2004, vol. 2, pp. 889–894.

[12] T. Kanade, J. F. Cohn, and Y. L. Tian, "Comprehensive database for facial expression analysis," in *Proc. 4th IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2000, pp. 46–53.

[13] Y. Liu, K. Schmidt, J. Cohn, and S. Mitra, "Facial asymmetry quantification for expression-invariant human identification," *Comput. Vision Image Understanding*, vol. 91, no. 1/2, pp. 138–159, 2003.

[14] M. H. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform," *IEEE Trans. Acoust., Speech Signal Process.*, vol. ASSP-30, no. 2, pp. 140–154, Apr. 1982.

[15] M. Savvides, B. V. K. V. Kumar, and P. Khosla, "Face verification using correlation filters," in *Proc. 3rd IEEE Automatic Identification Advanced Technologies*, Tarrytown, NY, 2002, pp. 56–61.

[16] M. Savvides and B. V. K. Kumar, "Eigenphases vs. eigenfaces," in *Proc. IEEE Int. Conf. Pattern Recognition*, Cambridge, U.K., Aug. 2004, vol. 3, pp. 810–813.

[17] M. Savvides, B. V. K. Kumar, and P. K. Khosla, "Corefaces-robust shift invariant PCA based correlation filter for illumination tolerant face recognition," in *Proc. Conf. Computer Vision and Pattern Recognition*, Washington, DC, Jun. 2004, vol. 2, pp. 834–841.

[18] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[19] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[20] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," in *Data Mining and Knowledge Discovery*, U. Fayyad, Ed. Norwell, MA: Kluwer, 1998, vol. 2, pp. 121–167.

[21] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed. New York: Wiley, 1984.

[22] X. Liu, T. Chen, and B. V. K. V. Kumar, "Face authentication for multiple subjects using eigenflow," *Pattern Recog., Special Issue Biometr.*, vol. 36, pp. 313–328, 2003.

[23] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Maui, HI, Jun. 1991, pp. 584–591.

[24] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Belmont, CA: Thomson, 2002.

Sinjini Mitra received the B.Sc. degree in statistics from Presidency College, Calcutta, India, the M.Stat. degree from the Indian Statistical Institute (ISI), Calcutta, and the Ph.D. degree in statistics from Carnegie Mellon University (CMU), Pittsburgh. PA.

Currently, she is a Research Staff Member with the Information Sciences Institute, University of Southern California (USC), Marina del Rey. Prior to joining USC, she was a Postdoctoral Research Associate with the Department of Statistics and CyLab at CMU and a Visiting Research Scientist with the Indian Statistical Institute in 2005. Her research interests include statistical pattern recognition, biometric authentication and security, machine learning, and statistical methodologies for data analysis and inference.

Dr. Mitra is a member of the American Statistical Association (ASA) and its section on Statistics in Defense and National Security (SDNS) and the Institute of Mathematical Statistics (IMS). Her publications include six book chapters (one invited and five refereed) and about 18 technical papers. She has been a Reviewer for several journals and conferences, including the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and the *Journal of the American Statistical Association* (JASA). She received a graduate student travel award from CMU in 2004–2005 and several awards for excellence during her undergraduate and graduate studies in Calcutta, India.

Marios Savvides (M'04) received the B.Eng. degree (Hons.) in microelectronics systems engineering from the University of Manchester Institute of Science and Technology (UMIST), Manchester, U.K., in 1997, the M.Sc. degree in robotics from the Robotics Institute in Carnegie Mellon University (CMU), Pittsburgh, PA, in 2000, and the Ph.D. degree in biometrics identification technology from the Electrical and Computer Engineering Department, CMU, in 2004.

Currently, he is a Research Assistant Professor with a joint appointment in the Electrical and Computer Engineering Department and CyLab, CMU. His Ph.D. dissertation "Reduced complexity face recognition using advanced correlation filters and fourier subspace methods for biometric applications" developed novel pattern-recognition methods that work extremely well for face recognition but also provided technology that is directly applicable to other biometric modalities, such as fingerprinting and irises. He then stayed on as a Systems Scientist from 2004 to 2005. His research interest is in biometric recognition of face, iris, and fingerprint and palmprint modalities. He has authored and coauthored many conference and journal articles in the biometrics area and has filed at least two patent applications in this field.

Dr. Savvides is a member of SPIE and serves on the Program Committee of several biometric conferences, such as the IEEE AutoID, SPIE Biometric Identification Technologies and International Conference on Image Analysis and Recognition (ICIAR). He is a Co-Chair Organizer of the special biometrics session of the International Workshop on Multimedia Content Representation, Classification and Security. He was one of invited panelists in SPIE's Defense and Security Biometrics Conference, discussing the current biometric barriers and how to break them. He is also listed in the 2006 Edition of *Marquis' Who's Who in America*.

B. V. K. Vijaya Kumar (SM'90) received the B.Tech. and M.Tech. degrees in electrical engineering from Indian Institute of Technology, Kanpur, India, and the Ph.D. degree in electrical engineering from Carnegie Mellon University (CMU), Pittsburgh, PA.

Currently, he is a Professor with the Department of Electrical and Computer Engineering (ECE) and CyLab, CMU, where he has been since 1982. He served as Associate Department Head of the ECE Department from 1994 to 1996 and as Acting Head of the Department from 2004 to 2005. His research interests include automatic target-recognition algorithms, biometric-recognition methods, and coding and signal processing for data-storage systems. Currently, he is leading coding and signal processing research efforts in the data-storage systems center (DSSC) and the biometrics research efforts in CyLab at CMU. His publications include the book: *Correlation Pattern Recognition* (coauthored by Dr. A. Mahalanobis and Dr. R. Juday, Cambridge University Press, 2005), seven book chapters, and many technical papers.

Prof. Kumar has served on many conference program committees and was a Co-General Chair of the 2004 Optical Data Storage conference and a Co-General Chair of the 2005 *IEEE AutoID Workshop*. He is a Fellow of The International Society of Optical Engineering (SPIE) and a Fellow of the Optical Society of America (OSA). He is listed in various professional Who's Whos including Marquis' *Who's Who in Science and Engineering*, *Who's Who in the World*, and in the *American Men and Women of Sciences*. In 2003, he received the Eta Kappa Nu award for Excellence in Teaching in the Electrical Computer Engineering Department at CMU and the Carnegie Institute of Technology's Dowd Fellowship for educational contributions. He served as a Pattern Recognition Topical Editor for the Information Processing division of *Applied Optics*. He guest-edited the September 1990 special issue of *Optical Engineering* devoted to *Optical Pattern Recognition*, the April 2001 issue of IEEE JOURNAL OF SELECTED AREAS OF COMMUNICATION (JSAC) devoted to *Signal Processing for High-Density Storage Channels*, and the January 2004 special issue of *IEEE Signal Processing* magazine devoted to iterative soft processing. He is an Associate Editor for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.