Largest Matching Areas for Illumination and Occlusion Robust Face Recognition

Largest Matching Areas for Illumination and Occlusion Robust Face Recognition

Niall McLaughlin, Ji Ming, Danny Crookes

***Abstract*—In this paper we introduce a novel approach to face recognition which simultaneously tackles three combined challenges: uneven illumination, partial occlusion, and limited training data. The new approach performs lighting normalization, occlusion de-emphasis and finally face recognition, based on finding the *largest* matching area (LMA) at each point on the face, as opposed to traditional *fixed*-size local area based approaches. Robustness is achieved with novel approaches for feature extraction, LMA-based face image comparison and unseen data modeling. On the extended YaleB and AR face databases for face identification, our method using only a single training image per-person, outperforms other methods using a single training image, and matches or exceeds methods which require multiple training images. On the LFW face verification database, our method outperforms comparable unsupervised methods. We also show that the new method performs competitively even when the training images are corrupted.**

I. Introduction

A. Literature Review

In many real world applications of face recognition, it is impractical or impossible to obtain more than one training image per-person. However, the problem of face recognition given variable lighting, partial occlusion and a single training image per person is challenging. Both illumination and partial occlusion may result in large changes in the feature representation of the appearance of a person, where the changes caused by these factors can sometimes be greater than the variation between images of the same person [1], [2]. Given limited training data, it can be difficult to separate illumination and occlusion effects from inter-personal variation. In this paper, we aim to tackle these three problems together. Given a single training image per person, we develop a new algorithm for correcting the effects of illumination while disregarding the effect of partial occlusion.

For recognition with illumination variation, two main categories of approaches have been tried: those that attempt to model the appearance variation caused by illumination, and those that attempt to build an invariant representation of the face. Many of the original approaches to illumination invariant recognition involved statistical analysis of the variation of face images under differing illumination. It has been shown that a face, viewed from a single direction but under varying illumination, can be represented using a small number of eigenfaces [3], [4]. A set of illumination basis vectors can be constructed for each person using several training images taken from the same direction but under differing illumination conditions [5], [6], [7], [8]. Another statistics-based method is the spherical harmonics representation, which represents the surface appearance of an object, ignoring cast shadows, in a 9D linear subspace [9], [10], [11]. These statistical methods require several face images per person captured under differing illumination conditions to build the illumination models.

Other methods seek an illumination-invariant representation which allows a single training image to be used for each person. The Retinex illumination model [12] explains the observed brightness at each point on the face in terms of the intrinsic reflectivity at that point, together with the magnitude and angle of the incident illumination. Given this model, it is commonly assumed that the low spatial frequency components of the face image represent the illumination information, while the high spatial frequency components represent the intrinsic face reflectivity, that should be recovered for recognition. Thus, self quotient imaging (SQI) [13], [14], [15], [16] was used to approximate the illumination using a smoothed version of the face image, which is subtracted from the logarithm of the original face image to give an invariant representation. Similar filtering can be carried out directly in the frequency domain to remove the low-frequency illumination components, using the Fourier [17] or DCT [18] transforms. Taking the opposite approach, the gradient domain reveals the high frequency variation of pixels relative to their neighbors, which is invariant to low frequency illumination variation [19], [20]. Several gradient-domain methods have been developed, such as edge maps [21], Gradientfaces [22], robust gradient features [23], and subspace learning using gradient orientations [24] which can jointly handle illumination variation and partial occlusion. It is also possible to decompose the face image into Gabor features and learn which orientations and scales are most discriminant [25]. Similarly, local binary patterns (LBP) describe the appearance of a local region relative to its center point, and are therefore invariant to monotonic illumination change within the local area [26], [27]. Some recent methods make use of the commonalities between all human faces. For instance to generate synthetic face images to improve recognition in challenging conditions [28]. Symmetric shape from shading models (SSFS) use a generic 3D head model to recover the shape and albedo of any face using a single image [29], [30], while 3D morphable head model based methods re-project a face image to a canonical representation before recognition [31], [32], [33].

Face recognition given partial occlusion is another challenging problem that has received much attention. By definition, partial occlusion affects only certain parts of the face, while

The authors are with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, U.K. (e-mail: nmclaughlin02@qub.ac.uk, j.ming@qub.ac.uk, d.crookes@qub.ac.uk).

leaving others unaffected. As a result, use of local face descriptors, which should be unaffected by occlusion outside their locality, has been a popular method for tackling this problem. In [34] the face image is divided into a number of local areas, and the statistical variability of the training data within each local area is modeled using a Gaussian function. The similarity between two images is then defined as the sum of Mahalanobis distances between all the testing and training image local descriptors. A similar approach is taken by [35] and [36], where the training examples are used to learn statistics for the appearances of each local face area. During recognition, weights are assigned to every local area of the testing image, proportional to each area's likelihood given the unoccluded training examples. Using examples of occluded and unoccluded face areas, it is possible to train a classifier to identify which local areas of a testing image face are occluded. Examples of several different methods for modeling the appearance of local face areas, and of performing classification, exist in the literature. For example, local non-negative matrix factorization is adopted by [37] to model the local areas of a face, and a nearest-neighbor classifier is then used to classify each local area as belonging to the target class or to the occlusion class. Similarly, in [38] and [39], examples of occluded images patches are used to train an SVM classifier to detect occluded areas of a testing face image. In [40], the KL-divergence between local-binary-pattern histogram features is used to identify occluded areas; a KL-divergence threshold is learned from training data to decide whether a pair of image patches are similar enough to be included during recognition. The approach in [41] attempts to select the unoccluded face areas for recognition, using a maximum posterior probability method. While the use of local face descriptors has been popular in the literature, occlusion-robust recognition can also be performed without explicitly dividing the face into local regions, for example, face recognition by sparse representation [42], [43], or by random sampling [44]. Although such methods can recognize a face under occlusion, the authors of [42] found that superior results could be obtained by modeling each local area of the face individually. In effect, use of local descriptors takes advantage of our prior knowledge that many real-world occlusions cover contiguous areas of the face (in contrast to salt-and-pepper type noise which could be viewed as a type of non-contiguous occlusion).

B. Contributions of this Work

In this paper, we present a novel method for face recognition that is robust to simultaneous realistic partial occlusion and illumination variation. This method can be used with only a single training image per person. Based on the Retinex illumination model [12], we introduce a new method of comparing the similarity of face images, that is based on finding the *largest* matching area (LMA) at each face location. In contrast to existing methods for illumination normalization and occlusion modeling based on *fixed*-size local image areas, as described above, image comparison based on the largest matching areas both improves the discrimination between different persons, and at the same time, optimizes the accuracy of illumination normalization. In the new LMA method, we model the unseen illumination variation and occlusion with a novel unseen-data likelihood. This is included in a maximum a posteriori (MAP) framework to capture the largest matching local image areas for illumination normalization, and to deemphasize occluded local image areas during recognition. Experiments have been conducted showing the improved performance of the new method compared with existing methods from the literature.

The remainder of the paper is organized as follows. In Section II, we describe a piecewise constant illumination model which forms the basis for this paper's key idea: the largest matching area (LMA) approach. In Section III, the features used to represent each local face image area, and a method for comparing features given limited training data, are introduced. In Section IV, we first describe the LMA approach for identifying the largest matching image areas for face recognition with robustness to illumination changes, and we then explain how this approach can be further extended to provide robustness to partial occlusion. In Section V, experiments are performed to evaluate the new LMA approach and to compare with results from the literature. Finally, conclusions are presented in Section VI.

II. Largest Matching Area using Piecewise-Constant Illumination Model

Given a face image, let $I(x, y)$ refer to the brightness of the pixel at location (x, y), which depends on both the face's intrinsic reflectivity $R(x, y)$ and the illumination $L(x, y)$. It has been observed [12] that in natural images the illumination varies slowly, i.e., $L(x, y)$ is present at low spatial frequencies, while the intrinsic reflectance information $R(x, y)$ is present at higher spatial frequencies. This observation can be used to construct a model to separate the intrinsic reflectance $R(x, y)$ from the illumination $L(x, y)$ based on their differences in spatial variation. Specifically, we assume that surrounding each location (x, y) there is a small area $\delta(x, y)$ in which $L(i, j)$, for $(i, j) \in \delta(x, y)$, is approximately constant and can be described by two constant components: the first component is additive, modeling the mean pixel brightness within the area $\delta(x, y)$, and the second component is multiplicative, modeling the variance of the pixel brightnesses over the area $\delta(x, y)$. Therefore, inside $\delta(x, y)$ the observed pixel brightness value $I(i, j)$ at location (i, j) can be expressed as

$$I(i,j) \approx \beta_{\delta(x,y)} R(i,j) + \alpha_{\delta(x,y)} \quad \forall (i,j) \in \delta(x,y) \qquad (1)$$

where $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$ represent the respective constant additive illumination bias and constant multiplicative illumination factor associated with the area $\delta(x, y)$ surrounding pixel location (x, y). Equation (1) represents a piecewise constant illumination model: at each location (x, y), the illumination will remain even over a contiguous area $\delta(x, y)$ of pixels surrounding (x, y), while the intrinsic reflectance may change on a pixel-to-pixel basis. The values of $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$ can vary across different areas $\delta(x, y)$ of the face, to model globally uneven illumination over the whole image.

In this paper, we study face image recognition assuming that the illumination condition in each testing image can be

modeled by (1). For simplicity, we further assume that the local evenly-lit areas $\delta(x, y)$ have a square geometry. But we do not assume specific knowledge about the size of $\delta(x, y)$ and the values of $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$. We propose a novel approach to estimating the size of $\delta(x, y)$ for each location on the face, for normalizing the respective $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$ to achieve illumination-invariant face recognition. To maximize the system's performance, we assume that, while we have uncontrolled illumination for the testing images, we have well-lit training images that can be described by an image-specific global constant illumination model. The illumination model for the training images has an analogous expression to (1) but where $\delta(x, y)$ now represents a whole image and $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$ are the image-dependent constants. Later in the paper we show that this assumption can be relaxed.

We use a simplified example to illustrate our idea. Assume $\delta(x, y)$ is a square area on the testing image, for which the model (1) holds, and hence the constant illumination factors, $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$, can be normalized using mean and variance normalization, for example. Also assume the illumination in the corresponding area of the training image is similarly normalized, allowing both areas to be directly compared. Selecting a fixed size for the area $\delta(x, y)$ runs the risk that the normalization process will corrupt the intrinsic reflectance information, as some areas are too large to contain even illumination, while other areas are too small to allow normalization to be performed effectively. Ideally, we wish to find the *largest* area $\delta(x, y)$ in which the testing image can be modeled using the constant lighting condition (1), as this is likely to achieve the most accurate recovery of the intrinsic reflectance information.

Seeking an optimal solution to this problem, we propose the *Largest Matching Area* (LMA) approach: at each location (x, y), we find the largest testing image area that can be matched at the corresponding location by the training image, with an objective function that favors constant illumination. Since larger image areas with even illumination can be normalized more accurately, and since larger, correctly-normalized image areas can be more discriminative (than a smaller area at the same location) for correct matching, estimating the largest $\delta(x, y)$ based on identifying the largest matching image areas with even illumination optimizes recognition accuracy. The above example can be extended to other lighting normalization techniques which remove the constant illumination factors $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$ in order to recover the intrinsic reflectance information, as will be demonstrated in this paper.

Interestingly, we find that the LMA approach can also be used to help address the problem of partial occlusion. When performing face recognition given partial occlusion, the matching scores from occluded areas should be excluded, as these areas do not contribute to discrimination between the faces of different individuals. We show that the problem of deemphasizing the matching scores from occluded areas can be embedded within the problem of finding the largest matching areas between the training and testing images, where areas mismatching due to occlusion are deemphasized from the calculation. Hence, to some extent, the LMA framework is capable of simultaneously tackling the problems of illumination variation and partial occlusion. We will address all the problems mentioned above, the LMA approach for illumination normalization, its extension to deemphasizing partial occlusion and their combination in Section IV. Before this, in the following section we first describe the features used in our research to represent the face images.

III. Feature Representation and Likelihood

In this section we explain the feature representation used to describe each image area. Note that in this paper we assume all training and testing images have already been aligned. In terms of notation, we use I to represent a testing image and use $I_{\delta(x,y)}$ to represent a square area $\delta(x, y)$ in the image containing the pixels surrounding location (x, y), which can be modeled by the piecewise constant lighting condition (1) with unknown illumination factors $\alpha_{\delta(x,y)}$ and $\beta_{\delta(x,y)}$. We use I^m to represent a training image from person m, which assumes a global constant illumination model with image-dependent illumination factors α^m and β^m, and we use $I^m_{\delta(x,y)}$ to represent the area in the image corresponding to $I_{\delta(x,y)}$. During the feature extraction stage, we normalize the illumination difference between $I_{\delta(x,y)}$ and $I^m_{\delta(x,y)}$ by removing their respective constant illumination factors. Since we do not assume specific knowledge about the size of each $\delta(x, y)$, we perform lighting normalization and feature calculation over a range of different sizes for $\delta(x, y)$, assuming that the largest $\delta(x, y)$ for which (1) holds is contained within this range. Finally, we select the optimal estimate for the largest $\delta(x, y)$ to form the matching score for recognition. The algorithm for deciding the optimal estimate will be detailed in the next section (see Section IV). In this section we describe the methods used for calculating normalized features and for feature comparison, assuming (1) applies to each local area $\delta(x, y)$ being considered.

A. Band-Pass Filtering for Preprocessing

As a preprocessing step, we use a band-pass filter to reduce low-frequency illumination information corresponding to $\alpha_{\delta(x,y)}$ in (1), while retaining the discriminative reflectance information. The band-pass filter is also intended to remove high-frequency noise, as well as to further normalize any residual random, high-frequency illumination factors which can not be accurately represented by the piecewise constant illumination assumption. The filter we adopt is the Difference of Gaussians (DoG) kernel [13], [14], which has previously been shown to be effective for the task of illumination-invariant face recognition [27], [45]. Specifically, we use a centered spatial DoG kernel that can be expressed as follows

$$DoG(x, y) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-(x-y)^2/2\sigma_1^2} - \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-(x-y)^2/2\sigma_2^2} \tag{2}$$

where the parameters σ_1^2 and σ_2^2 specify the variances of the Gaussian functions and hence the bandwidth of the filter. In our experiments, their values are fixed at $\sigma_1^2 = 1$ and $\sigma_2^2 = 2$. We apply this filter to all the training and testing images before any other operations are carried out.

B. Fourier Magnitude Spectra as Features for Image Area Representation

Small changes in an individual's facial expression and/or head pose mean that even given well aligned face images, corresponding local areas of pixels may not correspond to the same physical locations on the faces. To reduce the effect of these small misalignment errors on recognition, we use the 2D Fourier magnitude spectrum as the feature to represent each local image area. By taking the magnitude spectrum, we omit the phase information, allowing us to take advantage of the shift invariance of the Fourier magnitude representation to improve the robustness to small misalignment errors and small facial expression changes. Applying the 2D Fourier transform to an area $I_{\delta(x,y)}$ in the band-pass filtered testing image, and assuming the piecewise constant illumination model (1) holds for the area, the resultant testing image magnitude spectrum for the area may be expressed as

$$|\tilde{\mathcal{I}}_{\delta(x,y)}(u,v)| \simeq \beta_{\delta(x,y)}|\tilde{\mathcal{R}}_{\delta(x,y)}(u,v)| + \delta(u,v)\tilde{\alpha}_{\delta(x,y)} \quad (3)$$

where $|\tilde{\mathcal{I}}_{\delta(x,y)}(u,v)|$ and $|\tilde{\mathcal{R}}_{\delta(x,y)}(u,v)|$ represent the respective magnitude spectra of the band-pass filtered pixel values and intrinsic reflectivity of the area, $\tilde{\alpha}_{\delta(x,y)}$ denotes the *residual* constant lighting background, if any, of the area after band-pass filtering, and $\delta(u,v)$ is the Kronecker delta function. We can obtain the analogous expression to (3) for a corresponding training image area $I^m_{\delta(x,y)}$. Suppose we have

$$|\tilde{\mathcal{I}}^m_{\delta(x,y)}(u,v)| \simeq \beta^m|\tilde{\mathcal{R}}^m_{\delta(x,y)}(u,v)| + \delta(u,v)\tilde{\alpha}^m \quad (4)$$

where $|\tilde{\mathcal{I}}^m_{\delta(x,y)}(u,v)|$ and $|\tilde{\mathcal{R}}^m_{\delta(x,y)}(u,v)|$ represent the pixel and reflectivity magnitude spectra of the area $I^m_{\delta(x,y)}$ after band-pass filtering, and as in (3) the residual constant lighting background of the area after band-pass filtering is denoted by $\tilde{\alpha}^m$. To form the feature vector for $I_{\delta(x,y)}$ we take the magnitude spectral coefficients $|\tilde{\mathcal{I}}_{\delta(x,y)}(u,v)|$, discard the zeroth coefficient $|\tilde{\mathcal{I}}_{\delta(x,y)}(0,0)|$, and concatenate the remaining coefficients into a vector. The feature vector for $I^m_{\delta(x,y)}$ is obtained in the same way. As indicated in (3) and (4), if constant lighting is assumed in $I_{\delta(x,y)}$ and $I^m_{\delta(x,y)}$, ignoring the zeroth spectral coefficient is equivalent to normalizing the constant additive illumination biases $\tilde{\alpha}_{\delta(x,y)}$ and $\tilde{\alpha}^m$ from the feature vectors. We use vectors $\mathbf{s}_{\delta(x,y)}$ and $\mathbf{s}^m_{\delta(x,y)}$ to represent the respective feature vectors obtained for the testing and training image areas $I_{\delta(x,y)}$ and $I^m_{\delta(x,y)}$, for each area $\delta(x,y)$.

C. Cosine Similarity Based Likelihood for Comparison

We use cosine similarity to compare the feature vectors from each pair of corresponding local areas from the training and testing images. The cosine similarity of the pair of feature vectors $\mathbf{s}_{\delta(x,y)}$ and $\mathbf{s}^m_{\delta(x,y)}$ can be expressed as

$$CS(\mathbf{s}_{\delta(x,y)}, \mathbf{s}^m_{\delta(x,y)}) = \frac{\mathbf{s}_{\delta(x,y)} \cdot \mathbf{s}^m_{\delta(x,y)}}{||\mathbf{s}_{\delta(x,y)}||||\mathbf{s}^m_{\delta(x,y)}||} \quad (5)$$

where the dot operation corresponds to the inner product between two vectors, and $||\mathbf{s}||$ represents the norm of vector $\mathbf{s}$. Cosine similarity is invariant to constant multipliers applied to the vectors being compared. Assume that the additive illumination biases of $\mathbf{s}_{\delta(x,y)}$ and $\mathbf{s}^m_{\delta(x,y)}$ have been removed, by band-pass filtering and discarding the residual zeroth spectral coefficient as described above, then cosine similarity will further cancel their multiplicative lighting difference, i.e., $\beta_{\delta(x,y)}$ and β^m in (3) and (4). Hence we can obtain an illumination-invariant similarity score between the two areas, assuming constant illumination.

It should be noted that the above procedures for illumination-invariant feature calculation and comparison will work most effectively if the sizes of the evenly-lit image areas $\delta(x,y)$ being considered are large. Hence it can be understood intuitively that searching for the largest areas $\delta(x,y)$ in the testing image in which the model (1) holds will lead to improved accuracy of lighting normalization, and therefore improved accuracy for face image recognition. In the algorithm to be described in the next section, we will use a likelihood function to measure how well a given testing image vector $\mathbf{s}_{\delta(x,y)}$ is matched by a training image vector $\mathbf{s}^m_{\delta(x,y)}$, in order to search for the largest matching image areas with even illumination. Following the usual convention in statistics, we write the likelihood function as $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$. This likelihood function is based on the cosine similarity (5) and takes an exponential function form

$$p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)}) = M^{CS(\mathbf{s}_{\delta(x,y)}, \mathbf{s}^m_{\delta(x,y)})} \quad (6)$$

where $M > 1$ is the base number of the likelihood function. For magnitude-spectrum based feature vectors, the cosine similarity define in (5) takes values in the range $[0, 1]$. The likelihood function (6) exponentially expands the difference in cosine similarity between different pairs of feature vectors. In this paper we have tested a range of base numbers M including e, to be detailed later. The likelihood function defined above can be used when there is only a single training image for each person.

IV. Image Matching for Face Recognition

In this section, we first introduce the largest matching area (LMA) approach for identifying the largest matching image areas with even illumination. This leads to an algorithm for face recognition with robustness to illumination variation. We then further extend this approach to improve robustness to unknown partial occlusion.

A. The LMA Approach

For face recognition, we consider comparing a testing image I against each of the training images I^m, by comparing the feature vectors $\mathbf{s}_{\delta(x,y)}$ and $\mathbf{s}^m_{\delta(x,y)}$ from locations (x,y), where within the compared areas $\delta(x,y)$, the testing image is assumed to have a constant lighting condition (1). At each location (x,y), we aim to find the largest area $\delta(x,y)$ over which this constant illumination assumption holds, thereby increasing the accuracy and discrimination of the feature vectors $\mathbf{s}_{\delta(x,y)}$ and $\mathbf{s}^m_{\delta(x,y)}$ used to compare the images. We formulate the problem of estimating the largest $\delta(x,y)$ as a maximum *a posteriori* (MAP) problem, based on the likelihood function defined in (6).

Given a testing image area represented by feature vector $\mathbf{s}_{\delta(x,y)}$, and assuming an equal prior probability P for all possible matching image areas $\mathbf{s}'$, we define the posterior probability of matching with the training image area $\mathbf{s}^m_{\delta(x,y)}$ as

$$P(\mathbf{s}^m_{\delta(x,y)}|\mathbf{s}_{\delta(x,y)}) = \frac{p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})P}{\sum_{\text{All } \mathbf{s}'} p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}')P} \simeq \frac{p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})}{\sum_{m'} p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^{m'}_{\delta(x,y)}) + p(\mathbf{s}_{\delta(x,y)}|\phi)} \tag{7}$$

where $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$ is defined in (6), as the likelihood that a given testing image area $\mathbf{s}_{\delta(x,y)}$ is matched by the training image area $\mathbf{s}^m_{\delta(x,y)}$. This likelihood is "accurate" if the lighting condition in the corresponding testing image area can be modeled by (1). The denominator, the average likelihood of $\mathbf{s}_{\delta(x,y)}$, is approximated by a sum of two terms. The first term is the average likelihood of $\mathbf{s}_{\delta(x,y)}$ over all the corresponding training image areas, assuming that $\mathbf{s}_{\delta(x,y)}$ will be matched by at least one of these training areas. Again, this average likelihood is accurate only if the given testing image area can be modeled by (1). The second term, $p(\mathbf{s}_{\delta(x,y)}|\phi)$, tries to model the likelihood of the given testing image area when it violates the constant illumination assumption, which can happen, for example, when the area $\delta(x,y)$ is too large. This term is also intended to model any testing image areas affected by occlusion, or by a combination of illumination and occlusion, which will also be addressed in this paper. In any of these unseen-data conditions, the likelihoods in the first term and in the numerator, $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$, will become invalid.

In our studies, we model the unseen-data likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$ using a mixture model, with resemblance to a conventional Gaussian mixture model (GMM). Specifically, we use a large set of natural image samples gathered from the training images, plus simulated random noisy image patches, to model testing image areas with variable lighting conditions and occlusion. Each sample image is represented using an identical feature representation to that used to model the face images: band-pass filtered Fourier magnitude spectra with constant-lighting normalization, extracted at every image location, and over a range of scales. We then group all the feature vectors with the same scale into a feature vector set, to simulate the unseen testing image areas of the same scale. Denote by B_δ the feature vector set consisting of all the feature vectors of the sample image areas with scale δ. Given a testing image vector $\mathbf{s}_{\delta(x,y)}$, and assuming that it will match at least one of the feature vectors in B_δ, but we do not necessarily know which, it is suitable to model the likelihood of $\mathbf{s}_{\delta(x,y)}$ by using a mixture model over B_δ. Specifically, we calculate the likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$ over the k-nearest neighbor (k-NN) set within B_δ

$$p(\mathbf{s}_{\delta(x,y)}|\phi) = \frac{1}{k} \sum_{\mathbf{s} \in B_\delta(k)} p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}) \tag{8}$$

where $B_\delta(k) \subset B_\delta$ is the feature set containing the k closest feature vectors for the given testing feature vector $\mathbf{s}_{\delta(x,y)}$ measured in terms of the likelihood (6). Further details of the implementation of (8) for selecting the sample image set and the value of k will be discussed in Section V.

Testing image areas with uneven illumination (and hence distorted feature vectors) are likely to have low likelihoods $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$ given the correct training image areas, but not necessarily low likelihoods given the incorrect training image areas (which is why we obtain recognition errors). The presence of the unseen-data likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$ helps to reduce the posterior probability (7) should such an erroneous match happen. In the event of occlusion, where low likelihoods $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$ may result for all the training image areas, the presence of the unseen-data likelihood prevents the posterior probability from growing large, due to the tending-to-zero of both the numerator and denominator likelihoods, which would give a false indication of matching if not corrected. Alternatively, for the testing image area $\mathbf{s}_{\delta(x,y)}$ with even illumination and matching training image area $\mathbf{s}^m_{\delta(x,y)}$, we can assume that the corresponding likelihood $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$ is most-likely greater than the corresponding unseen-data likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$, because

$$p(\mathbf{s}_{\delta(x,y)}|\phi) \simeq \frac{1}{k} p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)}) \leq p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)}) \tag{9}$$

In (9) above, the first approximation is based on the assumption that the matching, i.e., most-likely, training image area $\mathbf{s}^m_{\delta(x,y)}$, is included in $B_\delta(k)$, and will therefore dominate the mixture-based likelihood (8). The above inequality (9) indicates that for correctly matching training and testing image areas, the effect of the unseen-data likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$ tends to be small and hence, a large posterior probability can be obtained from (7). Therefore, posterior probability (7) can be used to identify the most-likely matching training and testing images areas with even illumination. Given a training image area and testing image area with even illumination, a large posterior probability will be obtained; conversely, a small posterior probability may indicate a mismatched training area and/or uneven illumination in the given testing area, assuming even illumination in the training image.

At each location (x,y), we aim to find the largest area $\delta(x,y)$ over which the testing image can be modeled by (1). This problem may be solved by estimating the largest matching testing image area, over the training images, with an objective function that favors constant illumination. Since larger image areas are generally more discriminative and can be normalized more accurately assuming even illumination, estimation based on searching for the largest matching image areas with even illumination optimizes the accuracy. We now reveal another important property of the posterior probability that allows the estimation to be carried out, i.e., it favors the continuity of match, by giving larger posterior probabilities to larger matching image areas with even illumination. Assume that $\mathbf{s}_{\delta(x,y)}$ and $\mathbf{s}^m_{\delta(x,y)}$ are a pair of matching testing and training image areas, in terms of constant lighting and having the greatest likelihood, i.e., $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)}) \geq p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^{m'}_{\delta(x,y)})$ for any $m' \neq m$, and $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)}) \geq p(\mathbf{s}_{\delta(x,y)}|\phi)$. Furthermore, assume $\epsilon(x,y)$ is a smaller area with the same

origin as $\delta(x,y)$ and is contained within $\delta(x,y)$. Then the above property says that:

$$P(\mathbf{s}^m_{\epsilon(x,y)}|\mathbf{s}_{\epsilon(x,y)}) \leq P(\mathbf{s}^m_{\delta(x,y)}|\mathbf{s}_{\delta(x,y)}) \tag{10}$$

That is, the posterior probability increases as the size of the matching image areas with even illumination increases. For clarity of presentation, the proof of inequality (10) is included in Appendix A.

Therefore based on (10), for each training image I^m, at each location (x,y), we can obtain an estimate of the largest area $\delta(x,y)$ of the testing image that can be modeled using the constant illumination assumption, by maximizing the posterior probability $P(\mathbf{s}^m_{\delta(x,y)}|\mathbf{s}_{\delta(x,y)})$ over $\delta(x,y)$ that favors the continuity of matching image areas with constant illumination. Given the unknown but fixed lighting condition of the testing image, we assume that the optimal estimate of the largest $\delta(x,y)$, in terms of the maximum posterior probability, will be obtained on the training image with the largest matching area of reflectance due to the inequality (10). We express the estimate found over I^m as

$$\hat{\delta}^m(x,y) = \tag{11}$$
$$\arg\left\{ P(\mathbf{s}^m_{\hat{\delta}^m(x,y)}|\mathbf{s}_{\hat{\delta}^m(x,y)}) = \max_{\delta(x,y)} P(\mathbf{s}^m_{\delta(x,y)}|\mathbf{s}_{\delta(x,y)}) \right\}$$

In face recognition, we form the matching score for each training image I^m by using the corresponding posterior of the estimate, i.e., $P(\mathbf{s}^m_{\hat{\delta}^m(x,y)}|\mathbf{s}_{\hat{\delta}^m(x,y)})$. Specifically, for every given testing image I, we calculate the posteriors $P(\mathbf{s}^m_{\hat{\delta}^m(x,y)}|\mathbf{s}_{\hat{\delta}^m(x,y)})$ for each training image I^m of each person m at every location (x,y). Then, the overall score for training image I^m matching the given testing image I can be defined as

$$\Gamma(I^m, I) = \sum_{(x,y)} \ln P(\mathbf{s}^m_{\hat{\delta}^m(x,y)}|\mathbf{s}_{\hat{\delta}^m(x,y)}) \tag{12}$$

where the sum is over all locations (x,y) within the images being compared. Algorithm 1 outlines the method for solving the estimation problem (11) and for calculating the match score (12).

In the above description, we have made two assumptions. The first assumption is that the training images can be modeled by image-specific constant illumination. This assumption effectively means that the size of the largest matching areas (i.e., $\delta(x,y)$), subject to constant illumination, is decided by the testing image only. Thus it helps to maximize the size of the matching areas that can be found between the training and testing images to obtain greater discrimination. If we drop this assumption and allow the training images to have random, piecewise constant lighting as in the testing images, the new algorithm may still be applicable. But in this case it will focus on the largest matching areas $\delta(x,y)$ in which *both* the training image *and* the testing image can have constant illumination. Such areas may be smaller than with the corresponding constantly-lit training images. In our experiments, we will demonstrate the applicability of the new algorithm to this situation, i.e., given more adverse training conditions. Our second assumption is that each person m has only one training image I^m. However, this algorithm can be easily extended to accommodate multiple training images for each person, by calculating the sum in the denominator of (7) over all the training images of all the persons, and by calculating an average match score (12) for each person over all the training images for the person.

Data: Testing image I, training image I^m, unseen-data feature vector sets B_δ
Result: Posterior matching score for each training image
for *each location* (x,y) *in testing image* I **do**
  for *each area size* $\delta(x,y)$ **do**
    Calculate feature vector $\mathbf{s}_{\delta(x,y)}$ for testing image area $I_{\delta(x,y)}$
    Calculate unseen-data likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$ using (8)
    for *each training image* I^m *and corresponding area* $\mathbf{s}^m_{\delta(x,y)}$ **do**
      Calculate likelihood $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$ using (6)
    end
    for *each training image* I^m *and corresponding area* $\mathbf{s}^m_{\delta(x,y)}$ **do**
      Calculate posterior $P(\mathbf{s}^m_{\delta(x,y)}|\mathbf{s}_{\delta(x,y)})$ using (7)
      Record the maximum posterior $P(\mathbf{s}^m_{\hat{\delta}^m(x,y)}|\mathbf{s}_{\hat{\delta}^m(x,y)})$ over $\delta(x,y)$, i.e., (11)
    end
    for *each person* m **do**
      Add the above obtained $\ln P(\mathbf{s}^m_{\hat{\delta}^m(x,y)}|\mathbf{s}_{\hat{\delta}^m(x,y)})$ to the match score (12)
    end
  end
end

Algorithm 1: Algorithm for computing the posterior matching score for each training image.

B. Extension to Occlusion Robustness

In (12) we assume that the optimal local areas selected for comparison, in terms of the maximum posterior probability of matching, will contain valid face features. We can further extend this method to model random partial occlusion of the testing images, in which some local areas do not contain valid face features. We assume that an occluded testing image area $\mathbf{s}_{\delta(x,y)}$ can be modeled by the unseen-data likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$, leading to a low posterior probability $P(\mathbf{s}^m_{\delta(x,y)}|\mathbf{s}_{\delta(x,y)})$ for any training face image areas $\mathbf{s}^m_{\delta(x,y)}$. Using this assumption we can improve robustness to partial occlusion by deemphasizing local areas with low posterior probabilities from the overall score for each face. In other words, instead of using all local areas for recognition, we aim to choose only the *reliable* local areas (defined in terms of large posterior probabilities) for recognition. This will retain as much inter-personal discriminative information as possible while improving robustness to partial occlusion.

The problem of identifying all reliable local areas can be formulated as a higher-level MAP problem, similar to (7), which makes use of the previously calculated local area posterior matching scores. Specifically, given a testing image I, we obtain its maximum local area posteriors $P(\mathbf{s}^m_{\hat{\delta}^m(x,y)}|\mathbf{s}_{\hat{\delta}^m(x,y)})$ for each training image I^m, along with the unseen-data posteriors $P(\phi|\mathbf{s}_{\delta(x,y)})$, which are identical for all the training images. The unseen-data posteriors are calculated using (7) with the numerator likelihood replaced by the unseen-data likelihood $p(\mathbf{s}_{\delta(x,y)}|\phi)$. We sort these probabilities in descending order. Let $P(\mathbf{s}^m_{\hat{\delta}^m(x_j,y_j)}|\mathbf{s}_{\hat{\delta}^m(x_j,y_j)})$ denote the sorted posterior probabilities for training image I^m, with $j = 1, 2, ..., J$ denoting the index of the face area locations in sorted order, from the highest posterior probability to the lowest posterior probability (assuming a total of J areas). Similarly, let $P(\phi|\mathbf{s}_{\delta(x_l,y_l)})$ denote the sorted unseen-data posterior probabilities, indexed by $l = 1, 2, ..., J$, from the highest posterior-probability area to the lowest posterior-probability area. To select the optimal local areas for recognition, we formulate a posterior probability for each training image I^m, using the corresponding $P(\mathbf{s}^m_{\hat{\delta}^m(x_j,y_j)}|\mathbf{s}_{\hat{\delta}^m(x_j,y_j)})$ for I^m, as a function of the number of local areas with the highest posterior probabilities. This posterior probability can be expressed as

$$P(I^m|I,\mathcal{J}) = \frac{\prod_{j=1}^{\mathcal{J}} P(\mathbf{s}^m_{\hat{\delta}^m(x_j,y_j)}|\mathbf{s}_{\hat{\delta}^m(x_j,y_j)})}{\sum_{m'} \prod_{j'=1}^{\mathcal{J}} P(\mathbf{s}^{m'}_{\hat{\delta}^{m'}(x_{j'},y_{j'})}|\mathbf{s}_{\hat{\delta}^{m'}(x_{j'},y_{j'})}) + \prod_{l=1}^{\mathcal{J}} P(\phi|\mathbf{s}_{\delta(x_l,y_l)})} \tag{13}$$

where $1 \leq \mathcal{J} \leq J$ is the number of the highest-posterior areas used in forming the posterior probability $P(I^m|I,\mathcal{J})$. Thus, the new matching score for training image I^m, in place of (12), can be defined as the maximum $P(I^m|I,\mathcal{J})$ over $\mathcal{J}$, i.e.,

$$P(I^m|I,\hat{\mathcal{J}}^m) = \max_{\mathcal{J}} P(I^m|I,\mathcal{J}) \tag{14}$$

where $\hat{\mathcal{J}}^m$ is an estimate of the number of the optimal local areas for comparing the testing image I against the training image I^m. In a similar way to the proof of inequality (10), we can show that the posterior (13) favors continuity of matching: higher posteriors are obtained when more local areas are matched (in terms of higher local area posteriors than any of the competitor areas including the unseen-data). Therefore, given a testing image, we can assume that an optimal estimate of all reliable local areas will be obtained on the training image with the largest number of matching areas, due to the maximum posterior. Eq. (14) can be viewed as an extension of the problem (11). Through finding the largest matching areas between the training and testing images, (11) obtains an estimate of the optimal local areas for lighting normalization, while (14) obtains an estimate of the optimal combination of local areas for deemphasizing occlusion.

V. Experiments

A. Data, Implementation and Experimental Conditions

In our experiments, we tested the proposed LMA system's ability to cope with both illumination variation, combined illumination variation with partial occlusion, and limited training data. Face identification experiments were carried out using two databases: the extended-YaleB database and the AR database, and a face verification experiment was carried out using the LFW database.

The extended YaleB database [7] contains the frontal face images of 38 persons, each captured under 64 illumination conditions without occlusion. As per the standard testing protocol, the images were split into 5 subsets based on illumination angle: Subset 1 $(0^\circ - 12^\circ)$, Subset 2 $(13^\circ - 25^\circ)$, Subset 3 $(26^\circ - 50^\circ)$, Subset 4 $(51^\circ - 77^\circ)$, and Subset 5 $(78^\circ - 90^\circ)$. A single image evenly illuminated image from Subset 1 was used as the training image for each person, i.e., Condition P00A+00E+00, and testing was carried out using all other images. Example images from each illumination subset are shown in Fig. 1.

The AR face database [46] consists of the frontal face images of 100 persons, captured with varying combinations of illumination, partial occlusion and facial expressions. This database was captured over two sessions spaced two weeks apart, and hence contains a relatively large degree of intra-personal appearance variation. A single clean image from Session 1, Condition 1, with even illumination, neutral facial expression and no occlusion was used as the training image for each person. All tests were carried out using images from Session 2, Conditions 14 - 26, with varying facial expressions, occlusions and lighting. The pre-aligned version of the AR database was used for all experiments. Examples images are shown in Fig. 2.

The Labeled Faces in the Wild (LFW) database [47] consists of around 13,000 face images from 1680 individuals, where all the faces images have been collected from the Web. The face images therefore exhibit a high degree of variation in illumination, pose, and expression, the only restriction being that faces were detected using the Viola-Jones detector [48]. The deep-funneled version of LFW [49] was used in our experiments, and all images were cropped to include only the face area.

The LMA system's parameters were kept consistent across all databases. Specifically, all images were preprocessed using a DoG filter with parameters $\sigma_1 = 1$, $\sigma_2 = 2$, in (2), as in [27]. Optimization over the size of the local image areas $\delta(x, y)$, which had square geometry, in (11), was carried out over the following range of scales: 5×5, 7×7, 9×9, ..., 29×29 pixels. The features, representing local image areas at each scale, were sampled from a grid covering the whole face image with a stride of 5 pixels. The base number M in (6) was fixed to 1×10^9. The unseen-data likelihood model, $p(\mathbf{s}_{\delta(x,y)}|\phi)$ in (8), was built on a feature set B_δ consisting of the feature vectors of 1000 image areas at each scale δ, randomly selected from all training image areas of all persons. In our experiments, slightly better performance was obtained by augmenting this feature set with the feature vectors of randomly generated square image patches at each scale δ, where each pixel value was drawn from normal distribution $\sim \mathcal{N}(128, 16^2)$. Over this feature set, the k closest feature vectors were chosen for each testing image area at each scale to calculate $p(\mathbf{s}_{\delta(x,y)}|\phi)$. In our experiments, k was set to 10.

In Section V-F, we will investigate the sensitivity of the system to differing parameters.

For real-world applications efficiency may be important. Our system takes around 0.02 seconds to compare two face images on a single-core 3GHz machine in Matlab. The time needed to search a large gallery could be reduced as the task of comparing a probe with the gallery is inherently parallel. The majority of our system's execution time is used for feature extraction, however in practice this step only needs to be performed once per face image and the results stored for later use.

B. Recognition Results on Extended YaleB

We first compare the face recognition performance of the proposed LMA system on extended YaleB with results from the literature. The comparison is presented in Table I. It can be seen that for illumination Subsets 1 to 3, all systems were able to attain perfect accuracy. We observed that the illumination change from the training data in these subsets is not significant and does not pose an issue for recognition. Illumination Subsets 4 and 5 contain much larger variation from the training data. For Subset 4, our system produced results comparable with the existing literature, and for Subset 5, containing the most challenging illumination conditions, our proposed system was able to outperform ELT [27], PCML [50] and L&S [51], and achieved the same accuracy as OLHE [52] which used all images from Subset 1 for training. We also compare with IGO-LDA [24], which achieves an average accuracy over all subsets of 97.80%, compared with LMA's average accuracy of 99.62%. Table I also includes the LMA results without using the unseen-data model, which will be discussed later.

C. Recognition Results on AR

We now compare the face recognition performance of our LMA system tested on the AR database with results cited from the literature. Table II shows the comparison. When training the LMA system, a single image with frontal lighting from Session 1, Condition 1 was used for each person. The examples taken from the literature for comparison all used more than one training image per-person. Images from Session 2 were used during testing.

From Table II it can be seen that the LMA system is able to perform better than or comparably with existing systems from the literature [42], [53], [54], [55], particularly for occluded conditions, achieving 98% recognition rate with sunglasses occlusion, and 96% with scarf occlusion. These results are notable, as only a single training image per person was used by the LMA system, while the other systems being compared with used 8 training images per person. On the clean conditions the LMA system achieved an accuracy rate of 99% (the only error occurred as a result of a corrupted testing image W-027-14.bmp).

D. Fixed Size versus Maximum Size

The LMA system searches for the largest matching area with constant illumination between the testing and training images at each point on the face. In this experiment, we compare fixed-size area based recognition with the LMA approach. We use the same posterior probability based scoring algorithm (12), with the optimal estimate of the matching area $\hat{\delta}^m(x,y)$ replaced by fixed-size area $\delta^m(x,y)$ for all (x,y) and m. We have tested a range of scales for the fixed $\delta^m(x,y)$. The comparison is intended to confirm our intuition that by maximizing the size of matching area at each point of the face, discrimination and accuracy for lighting normalization improves over the use of fixed-size local areas.

The accuracy rates obtained for each fixed local area size from 5×5 up to 29×29 are shown in Fig. 3 for the extended YaleB database and in Fig. 4 for the AR database. In addition the recognition accuracy produced by the LMA system is included in both figures as a horizontal line for comparison. In both figures, for recognition based on fixed-size local areas, as the size increases, recognition accuracy also increases until a certain size. For very small areas it is unlikely that there will be enough discriminative information to accurately match between testing and training images. Also for very small areas the lighting normalization process may cause corruption to the intrinsic reflectance of the image, further degrading the discrimination. Conversely for very large areas, illumination variation, occlusion and other factors can again make accurate normalization difficult, even between images of the same individual. Ideally, we would like to base recognition on the maximum local area size at each location, as this is likely to achieve the best accuracy for normalization and discrimination. However, it is not always possible to know the optimal local area size *a priori*, as it depends on the characteristics of the given testing data and training data. By comparing the recognition accuracy curve in Fig. 3 for the extended YaleB database with that of Fig. 4 for the AR database, it can be seen that optimal recognition accuracy occurs at different local area sizes in both databases. The proposed LMA system exceeds, in the case of the AR database, and is comparable with, in the case of the extended YaleB database, the best accuracy with fixed-size local areas. Hence the LMA approach is capable of automatically adapting to different databases and testing conditions.

E. Unseen-Data Model

In this analysis, we examine the ability of the LMA system to cope with training and testing mismatch that could be caused by illumination variation, partial occlusion, or other types of variation. In particular, we investigate the effectiveness of modeling the mismatch by using the unseen-data likelihood, i.e., $p(\mathbf{s}_{\delta(x,y)}|\phi)$, in the posterior probability (7) for improving recognition accuracy. The comparison was conducted between the LMA systems with/without including this model. The results for both the extended YaleB and AR databases by the LMA system without this model are included in Table I and Table II, respectively.

In the YaleB database, which does not include any realistic partial occlusion, use of the unseen-data model improves recognition accuracy for Subset 5, which contains the most challenging illumination conditions. This may be due to very

dark areas of the face being captured by the unseen-data model and hence being deemphasized from the posterior probability of the whole face. These very dark areas are more likely to be corrupted by noise and hence less likely to contribute discriminative information. In the AR database, including the unseen-data model also leads to an improvement in recognition accuracy for the majority of testing conditions. The improvement is more significant for some severe testing conditions, such as screaming which causes a large expression change, and combined occlusion/illumination variation (e.g., sunglasses + illumination changes). In two cases, we observed a small drop in accuracy.

A new set of experiments was conducted, to systematically assess the robustness of the LMA system to varying amounts of partial occlusion. The extended YaleB database was used in the new experiments, to simulate testing conditions containing both varying amount of occlusion and at the same time significant illumination variation. For each person, a single un-occluded, evenly-illuminated face image was used as the training image. During testing, images with simulated partial occlusion were generated by randomly replacing a contiguous square of the testing image, with an unrelated image, covering 20%, 40%, 60% or 80% of the testing image area. During testing, no knowledge of the position and nature of the occlusion was provided to the system. Example testing images with simulated occlusion are shown in Fig. 5. In this experiment, for each occlusion amount, the average recognition accuracy over all illumination subsets was recorded. Similar experiments were conducted previously by other researchers (e.g., [56] and [42]). Our LMA-based results, with/without including the unseen-data model, are shown in Figure 7 along with the results from the literature. As expected, when the amount of occlusion is increased the performance of the system decreases. However, use of the unseen-data model in the LMA system significantly increases recognition accuracy compared to without this model. The results produced by the LMA system compare favorably with those of [56] and [42]. The posterior probability of a match between a clean face image and realistic images with: facial expression change, illumination change, and partially occluded local areas, is visualized in Fig. 6.

F. Parameter Sensitivity

The sensitivity of the LMA system to varying the values of parameters was tested, particularly, the base number M used in the likelihood function (6), and the k-NN set cardinality k used in calculating the unseen-data likelihood (8). This experiment was performed using the AR database, and while one parameter was varied, the other parameters remained at the values described at the beginning of this section.

The change in recognition accuracy that results from varying the value of the base number M is shown in Fig. 8. As the value of M is increased from e to 10^9, the overall recognition accuracy also slightly increases, however the performance of the system is not highly sensitive to the specific value of this parameter used. As mentioned earlier, a value of $M = 10^9$ was used to obtain the results presented above. The change in recognition accuracy that results from varying the value of the k parameter is shown in Fig. 9. It can be seen that when the k parameter is decreased from 1000 to 10, the recognition accuracy increases. Small values of k may lead a more accurate estimate of the likelihood of a given feature vector. However, the system is again not highly sensitive to the specific value of this parameter used.

G. Misalignment Robustness

When performing face recognition, the probe image must be accurately aligned with the gallery images to allow corresponding face regions to be compared. In this experiment we investigate the robustness of magnitude Fourier features to small misalignment errors, compared with discrete cosine transform (DCT) features. This experiment was carried out using the extended-YaleB database, and identification accuracy was averaged over all illumination subsets. During testing, all probe images were translated diagonally by a specified number of pixels, simulating misalignment of the probe and gallery images.

The results presented in Table III show that magnitude Fourier features are more robust to misalignment error than DCT features. Magnitude Fourier features even gives some benefit when no misalignment is introduced, suggesting that these features may be more robust in the general case.

H. Different Training Images

In many situations it may not be possible to obtain an un-occluded, evenly illuminated, training image for each person. As mentioned in Section IV-A, the LMA approach can be applied in situations where both the training and testing data have varying lighting conditions that can be modeled by (1). The AR database was used to demonstrate this. In contrast to the earlier experiments where a clean image was used for training, in this experiment, corrupted images were used for training, including those with partial occlusion, illumination variation, or combinations of both corruption conditions, with the same testing conditions.

Table IV shows the results, where for each training condition, the recognition accuracy results were averaged over all testing conditions. We can see that the best overall average recognition accuracy was achieved with the training image without occlusion and with the right illumination condition, although this accuracy is only marginally better than that achieved using the non-occluded, frontal-illumination condition. The lowest overall performance was observed when the scream condition was used as the training image for each person. As the LMA system was not specifically designed to handle expression changes, this is not unexpected. Again, we see the importance of the unseen-data model for improving robustness, compared to the system without this model.

I. Face Verification

We assess the face verification performance of the proposed system using the Labeled Faces in the Wild (LFW) database [47]. In contrast with recent approaches making use

of deep neural networks [57] which require large training sets, our face verification system is *unsupervised*, meaning we do not make use of the LFW training set for setting the system's parameters to better distinguish between same and different face image pairs. Therefore the system's parameters remain identical to those used in all previous experiments. For given pair of images, we calculate the posterior probability of matching, using a modified version of (7), where the denominator is calculated using the union of a large set of face images randomly selected from the LFW training-set, and one of the images in the verification pair. And the unseen data likelihood, (8), is calculated using a large set of randomly selected image patches from the LFW training-set.

The standard LFW testing protocol was followed, with results calculated using 10-fold cross-validation using the suggested data splits. Overall results are presented as an ROC curve, see Fig. 10. Our proposed LMA system achieved an Area Under the Curve (AUC) score of 0.8304, which is superior to other unsupervised face verification methods such as, Local higher-order statistics [58] with an AUC of 0.8107, or, Locally Adaptive Regression Kernels [59] with an AUC of 0.7830. Although some other unsupervised methods achieve better performance (e.g. [60] AUC of 0.9405), they make use of additional pre-processing steps such as fitting 3D head models to cope with pose variability [60]. We showed in our method that some robustness to pose variation may also be achieved by finding the largest matching areas between the pose varying images for comparison.

VI. Conclusion

In this paper we have presented a novel approach to achieving robustness in face recognition, based on finding the largest matching area at each point. Our method tackles three combined challenges simultaneously: uneven illumination, partial occlusion, and having only a single training image per person. Experiments show that, compared with other methods from the literature, our method outperforms methods which use only a single training image per person, and matches or outperforms methods which require multiple training images. Further tests in the paper show the importance of two factors in the success of the new method: the inclusion of an *unseen-data model*, and finding the (dynamic) largest matching area rather than a fixed-size patch. It is also shown that performance is much better if the single training image is good quality and non-occluded, although the method still performs competitively when the training image is corrupted.

Appendix A: Proof of Inequality (10)

This inequality can be proven if we can assume that, at a given location, larger areas will normally be more discriminative than smaller areas to identify if the two areas are matching or not. This can be expressed in terms of likelihood ratios:

$$\frac{p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})}{p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^{m'}_{\delta(x,y)})} \geq \frac{p(\mathbf{s}_{\epsilon(x,y)}|\mathbf{s}^m_{\epsilon(x,y)})}{p(\mathbf{s}_{\epsilon(x,y)}|\mathbf{s}^{m'}_{\epsilon(x,y)})} \tag{15}$$

where, as defined in (10), $\mathbf{s}_{\delta(x,y)}$ and $\mathbf{s}^m_{\delta(x,y)}$ are two matching feature vectors in terms of having the greatest likelihood ratio $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})/p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^{m'}_{\delta(x,y)})$ for all $\mathbf{s}^{m'}_{\delta(x,y)} \neq \mathbf{s}^m_{\delta(x,y)}$, and $\mathbf{s}_{\epsilon(x,y)}$ is a sub-vector in $\mathbf{s}^m_{\delta(x,y)}$ representing a smaller image area, with $\mathbf{s}^m_{\epsilon(x,y)}$ representing the corresponding matching sub-vector in $\mathbf{s}^m_{\delta(x,y)}$, and $\mathbf{s}^{m'}_{\epsilon(x,y)}$ representing the corresponding mismatching sub-vector in $\mathbf{s}^{m'}_{\delta(x,y)}$. Equation (15) simply indicates that larger likelihood ratios will be obtained based on larger image areas to differentiate between matching and mismatching images at the given location. Based on (9), we can have a similar inequality concerning the likelihood ratio associated with the unseen data:

$$\frac{p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})}{p(\mathbf{s}_{\delta(x,y)}|\phi)} \geq \frac{p(\mathbf{s}_{\epsilon(x,y)}|\mathbf{s}^m_{\epsilon(x,y)})}{p(\mathbf{s}_{\epsilon(x,y)}|\phi)} \tag{16}$$

Dividing both the numerator and denominator of (7) by $p(\mathbf{s}_{\delta(x,y)}|\mathbf{s}^m_{\delta(x,y)})$, and applying the above two likelihood-ratio inequalities to the expression, we can obtain the posterior probability inequality (10).

References

[1] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 721–732, Jul. 1997.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, pp. 399–458, Dec. 2003.

[3] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, Jun. 1994, pp. 995–999.

[4] A. Shashua, "On photometric issues in 3d visual recognition from a single 2d image," *International Journal of Computer Vision*, vol. 21, pp. 99–122, 1997.

[5] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[6] A. Batur and M. Hayes, "Linear subspaces for illumination robust face recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, Kauai, HI, USA, 2001, pp. 296–301.

[7] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, Jun. 2001.

[8] P. Belhumeur and D. Kriegman, "What is the set of images of an object under all possible lighting conditions?" in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, Jun. 1996, pp. 270–277.

[9] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object," *Journal of the Optical Society of America*, vol. 18, no. 10, pp. 2448–2459, Oct. 2001.

[10] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 218–233, 2003.

[11] L. Qing, S. Shan, and X. Chen, "Face relighting for face recognition under generic illumination," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, no. 5, Montreal, Quebec, Canada, May 2004, pp. 733–736.

[12] H. Wang, S. Li, and Y. Wang, "Face recognition under varying lighting conditions using self quotient image," in *IEEE International Conference on Automatic Face and Gesture Recognition*, May 2004, pp. 819–824.

[13] ——, "Face recognition under varying lighting conditions using self quotient image," in *IEEE Conference on Automatic Face and Gesture Recognition*, Seoul, Korea, 2004, pp. 819–824.

[14] ——, "Generalized quotient image," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, Washington, DC, USA, Jun. 2004, pp. 498–505.

[15] V. Štruc and N. Pavešić, "Illumination invariant face recognition by non-local smoothing," in *Conference on Biometric ID management and multimodal communication*, Madrid, Spain, 2009, pp. 1–8.

[16] S. Srisuk and A. Petpon, "A gabor quotient image for face recognition under varying illumination," in *Advances in Visual Computing, 4th International Symposium*, Las Vegas, NV, USA, 2008, pp. 511–520.

[17] S.-I. Choi and G.-M. Jeong, "Shadow compensation using fourier analysis with application to face recognition," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 23–26, 2011.

[18] V. P. Vishwakarma, S. Pandey, and M. N. Gupta, "An illumination invariant accurate face recognition with down scaling of dct coefficients," *Journal of Computing and Information Technology*, pp. 53–67, 2010.

[19] S.-D. Wei and S.-H. Lai, "Robust face recognition under lighting variations," in *International Conference on Pattern Recognition*, vol. 1, Istanbul, Turkey, Aug. 2004, pp. 354–357.

[20] C.-H. T. Yang, S.-H. Lai, and L.-W. Chang, "Robust face matching under different lighting conditions," vol. 2, Lusanne, Switzerland, 2002, pp. 149–152.

[21] Y. Gao and M. Leung, "Face recognition using line edge map," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 764–779, Jun. 2002.

[22] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2599–2606, 2009.

[23] C.-X. Ren, Z. Lei, D.-Q. Dai, and S. Li, "Enhanced local gradient order features and discriminant analysis for face recognition," *Cybernetics, IEEE Transactions on*, vol. PP, no. 99, pp. 1–14, 2015.

[24] G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "Subspace learning from image gradient orientations," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 12, pp. 2454–2466, Dec 2012.

[25] C.-X. Ren, D.-Q. Dai, X.-X. Li, and Z.-R. Lai, "Band-reweighed gabor kernel embedding for face image representation and recognition," *Image Processing, IEEE Transactions on*, vol. 23, no. 2, pp. 725–740, 2014.

[26] G. Heusch, Y. Rodriguez, and S. Marcel, "Local binary patterns as an image preprocessing for face authentication," in *International Conference on Automatic Face and Gesture Recognition*, Southampton, UK, Apr. 2006, pp. 6–14.

[27] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, 2010.

[28] Y. Xu, X. Fang, X. Li, J. Yang, J. You, H. Liu, and S. Teng, "Data uncertainty in face recognition," *Cybernetics, IEEE Transactions on*, vol. 44, no. 10, pp. 1950–1961, Oct 2014.

[29] W. Y. Zhao and R. Chellappa, "Illumination-insensitive face recognition using symmetric shape-from-shading," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, Hilton Head, SC, USA, 2000, pp. 286–293.

[30] T. Sim and T. Kanade, "Combining models and exemplars for face recognition: An illuminating example," in *Workshop on Models versus Exemplars in Computer Vision, IEEE Computer Vision and Pattern Recognition*, Kauai, HI, USA, Dec. 2001.

[31] V. Blanz and T. Vetter, "Face recognition based on fitting a 3d morphable model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063–1074, Sept. 2003.

[32] Y. Wang, L. Zhang, Z. Liu, G. Hua, Z. Wen, Z. Zhang, and D. Samaras, "Face relighting from a single image under arbitrary unknown lighting conditions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1968–1984, Nov. 2009.

[33] I. Kemelmacher-Shlizerman and R. Basri, "3d face reconstruction from a single image using a single reference face shape," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 394–405, Feb. 2011.

[34] A. M. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 748–763, Jun. 2002.

[35] J. Seo and H. Park, "A robust face recognition through statistical learning of local features," in *Neural Information Processing*, ser. Lecture Notes in Computer Science, B.-L. Lu, L. Zhang, and J. Kwok, Eds., 2011, vol. 7063, pp. 335–341.

[36] ——, "Robust recognition of face with partial variations using local features and statistical learning," *Neurocomputing*, vol. 129, pp. 41–48, 2014.

[37] H. J. Oh, K. M. Lee, and S. U. Lee, "Occlusion invariant face recognition using selective local non-negative matrix factorization basis images," *Image and Vision Computing*, vol. 26, no. 11, pp. 1515–1523, 2008.

[38] Z. Chen, T. Xu, and Z. Han, "Occluded face recognition based on the improved svm and block weighted lbp," in *International Conference on Image Analysis and Signal Processing*, Oct. 2011, pp. 118–122.

[39] R. Min, A. Hadid, and J. Dugelay, "Improving the recognition of faces occluded by facial accessories," in *IEEE International Conference on Automatic Face Gesture Recognition and Workshops*, Mar. 2011, pp. 442–447.

[40] W. Zhang, S. Shan, X. Chen, and W. Gao, "Local gabor binary patterns based on kullback leibler divergence for partially occluded face recognition," *IEEE Signal Processing Letters*, vol. 14, no. 11, pp. 875–878, Nov. 2007.

[41] J. Lin, J. Ming, and D. Crookes, "Robust face recognition with partially occluded images based on a single or a small number of training samples," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2009, pp. 881–884.

[42] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[43] J. Wang, C. Lu, M. Wang, P. Li, S. Yan, and X. Hu, "Robust face recognition via adaptive sparse representation," *Cybernetics, IEEE Transactions on*, vol. 44, no. 12, pp. 2368–2378, Dec 2014.

[44] S. Fidler, D. Skocaj, and A. Leonardis, "Combining reconstructive and discriminative subspace methods for robust classification and regression by subsampling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 3, pp. 337–350, Mar. 2006.

[45] H. Han, S. Shan, X. Chen, and W. Gao, "A comparative study on illumination preprocessing in face recognition," *Pattern Recognition*, vol. 46, no. 6, pp. 1691 – 1699, 2013.

[46] A. Martinez and R. Benavente, "The AR face database, CVC technical report 24," 1998.

[47] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, October 2007.

[48] P. A. Viola and M. J. Jones, "Robust real-time face detection," in *IEEE International Conference on Computer Vision*, Vancouver, Canada, 2001, p. 747.

[49] G. Huang, M. Mattar, H. Lee, and E. G. Learned-Miller, "Learning to align from scratch," in *Advances in Neural Information Processing Systems*, 2012, pp. 764–772.

[50] N. McLaughlin, J. Ming, and D. Crookes, "Illumination invariant facial recognition using a piecewise-constant lighting model," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, Mar. 2012, pp. 1537–1540.

[51] X. Xie, W.-S. Zheng, J. Lai, P. C. Yuen, and C. Y. Suen, "Normalization of face illumination based on large-and small-scale features," *Image Processing, IEEE Transactions on*, vol. 20, no. 7, pp. 1807–1821, 2011.

[52] P. Lee, S.-W. Wu, and Y.-P. Hung, "Illumination compensation using oriented local histogram equalization and its application to face recognition," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4280–4289, Sept. 2012.

[53] X. Wei, C.-T. Li, and Y. Hu, "Face recognition with occlusion using dynamic image-to-class warping (dicw)," in *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*, April 2013, pp. 1–6.

[54] Z. Zhou, A. Wagner, H. Mobahi, J. Wright, and Y. Ma, "Face recognition with contiguous occlusion using markov random fields," in *Computer Vision, 2009 IEEE 12th International Conference on*, Sept 2009, pp. 1050–1057.

[55] L. Zhang, M. Yang, X. Feng, Y. Ma, and D. Zhang, "Collaborative representation based classification for face recognition," *arXiv preprint arXiv:1204.2358*, 2012.

[56] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: A simple deep learning baseline for image classification?" 2014, preprint http://arxiv.org/abs/1404.3606.

[57] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," *arXiv preprint arXiv:1406.4773*, 2014.

[58] G. Sharma, S. ul Hussain, and F. Jurie, "Local higher-order statistics (lhs) for texture categorization and facial analysis," in *European Conference on Computer Vision*, 2012, pp. 1–12.

[59] H. J. Seo and P. Milanfar, "Face verification using the lark representation," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1275–1286, Dec 2011.

[60] D. Yi, Z. Lei, and S. Z. Li, "Towards pose robust face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2013, pp. 3539–3545.

TABLE I

*Recognition accuracy, on the extended-YaleB database, comparing the LMA approach with the literature. The results show the number of correctly recognized images for each testing condition. The total number of testing images in each subset is shown in brackets in the Illumination Subset column. The symbol * indicates that more than one training image per person was used. We also show the results for our proposed approach without the unseen data model (LMA-UDM).*

Illumination Subset	LMA	LMA -UDM	ELT [27]	PCLM [50]	OLHE* [52]	L&S [51]
1 (263)	263	263	263	263	-	263
2 (456)	456	456	456	456	456	456
3 (455)	455	455	455	455	455	440
4 (526)	524	523	523	426	512	223
5 (714)	707	703	694	693	707	493

TABLE II

Recognition accuracy (%) on the AR database, comparing the LMA based approaches with the literature. For LMA, a single clean image from Session 1 was used for training and all images from Session 2 were used during testing. All the other systems for comparison use more than one training image per-person.

	Recognition Accuracy (%)					
Test Condition	LMA	LMA -UDM	SRC [42]	DICW [53]	MRF [54]	CRC [55]
Clean	99	99	-	-	-	-
Smile	96	94	-	-	-	-
Anger	99	99	-	-	-	-
Scream	66	56	-	-	-	-
Illum. Right	100	100	-	-	-	-
Illum. Left	100	100	-	-	-	-
Illum. Both	97	100	-	-	-	-
Sunglasses	98	96	97.5	99.5	99	94.2
Sg.+Illum. R	98	80	-	-	-	-
Sg.+Illum. L	93	80	-	-	-	-
Scarf	97	98	93.5	98	97.5	95.8
Scarf+Illum. R	95	94	-	-	-	-
Scarf+Illum. L	89	88	-	-	-	-

TABLE III

Identification accuracy, averaged over all illumination subsets of the extended-YaleB database. These results compare the robustness to misalignment of the proposed system with either 2D magnitude Fourier features or 2D DCT features.

Offset (pixels)	0	1	2	3	4	5	6
Mag. Fourier	99.91	99.91	99.74	96.5	89.55	78.94	60.95
DCT	97.07	97.06	95.18	89.93	77.12	62.77	38.38

Fig. 1. Example images from the extended YaleB database showing the typical illumination in each subset. A single image from illumination subset 1 was used as the training image for each person, and all other images were used during testing.

TABLE IV

Recognition accuracy (%) on the AR database, with a single corrupted training image from varying conditions, averaged over all testing conditions. Results are shown for the LMA system tested with and without the unseen-data model.

	Average Recognition Accuracy(%)	
Training Condition	LMA	LMA (no unseen-data model)
Clean	94.38	91.08
Smile	83.53	67.92
Anger	84.00	77.76
Scream	50.53	32.53
Illum. Right	94.53	86.69
Illum. Left	92.69	87.30
Illum. Both	89.53	83.84
Sunglasses	80.38	71.08
Sunglasses + Illum. Right	77.92	64.61
Sunglasses + Illum. Left	75.84	64.00
Scarf	81.23	70.69
Scarf + Illum. Left	76.84	62.69
Scarf + Illum. Left	74.53	63.53

Fig. 2. Example images from the AR database. In our experiments, either a single clean image or an imperfect image with nonuniform illumination and/or partial occlusion was used as training for each person, and all other images were used during testing. Testing and training images were selected from different sessions.

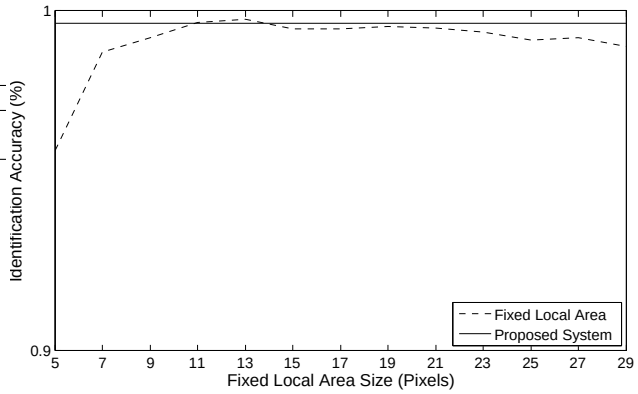

Fig. 3. Average recognition accuracy (%) on the extended YaleB database, for the system using fixed-size local areas as the size was varied. Also shown, as a solid line, is the recognition accuracy produced by the LMA system.

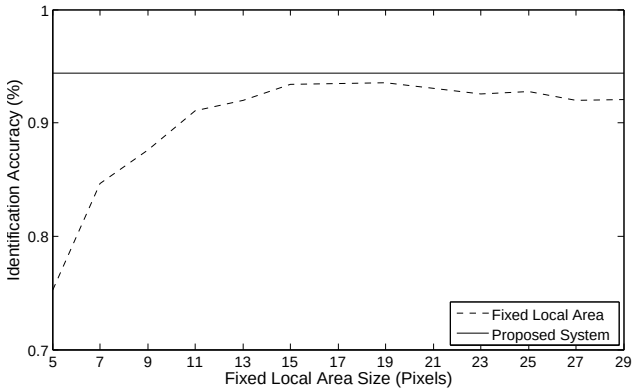

Fig. 4. Average recognition accuracy (%) on the AR database, for the system using fixed-size local areas as the size was varied. Also shown, as a solid line, is the recognition accuracy produced by the LMA system.

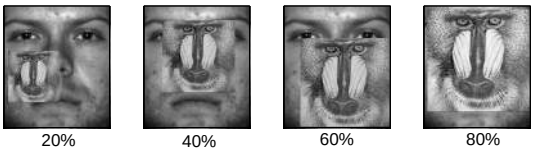

Fig. 5. Test images showing artificial occlusion, generated using the extended YaleB database. The percentage of the total image area covered by the occlusion is shown above each image.

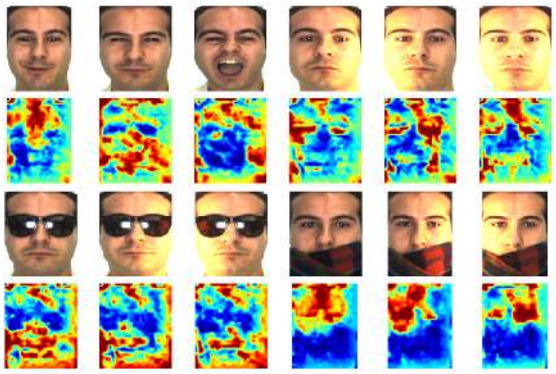

Fig. 6. Best viewed in color. This figure illustrates the posterior probability of a match between a clean training image and realistically corrupted testing images, where red indicates a high posterior probability, and blue indicates a low posterior probability. On the top row, the three images on the left show expression variation, and the three images on the right show illumination variation. The images on the bottom row show combined partial occlusion with illumination variation.

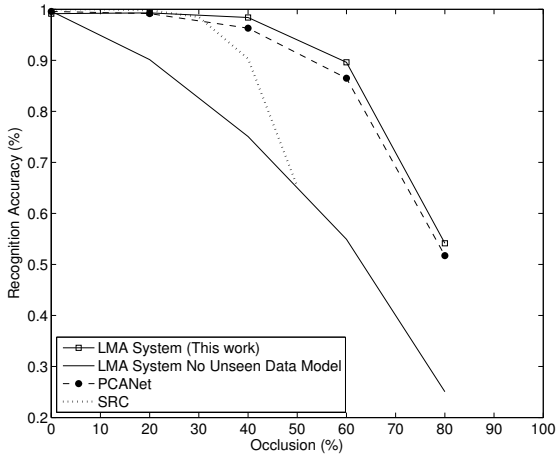

Fig. 7. Average recognition accuracy (%) over all testing subsets of the extended YaleB database as the percentage of artificial occlusion was varied. The LMA systems is compared to the PCANet [56] and SRC [42] systems.

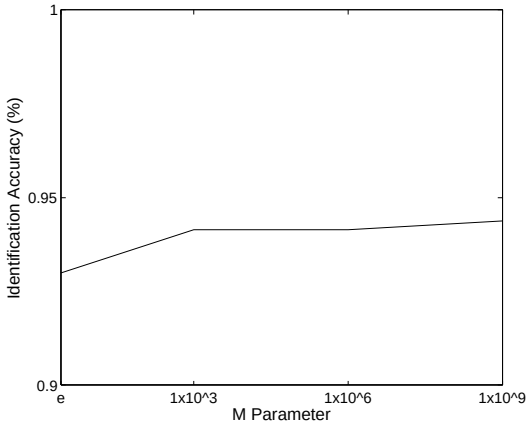

Fig. 8. Average recognition accuracy (%) over all testing subsets of the AR database, as the base number M in (6) was varied.

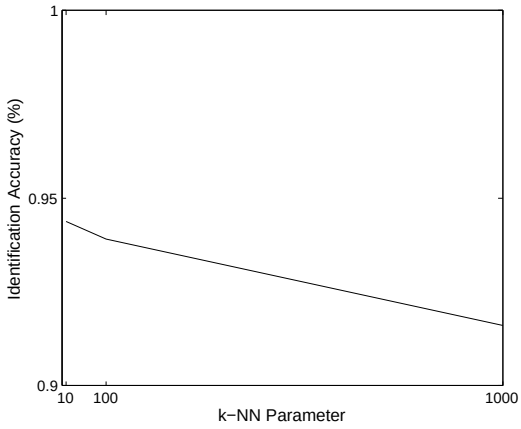

Fig. 9. Average recognition accuracy (%) over all testing subsets of the AR database, as the value of the k parameter in the k-NN unseen-data model (8) was varied.

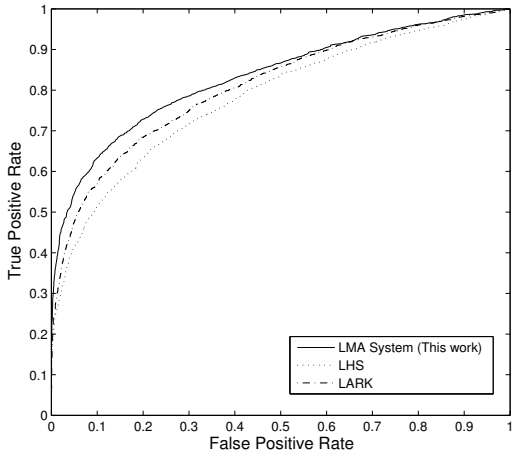

Fig. 10. ROC curve comparing performance on the LFW database of the LMA system with the LARK [59] and LHS [58] methods.