Illumination invariant face recognition and impostor rejection using different MINACE filter algorithms

Rohit Patnaik and David Casasent

Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

ABSTRACT

A face recognition system that functions in the presence of illumination variations is presented. It is based on the minimum noise and correlation energy (MINACE) filter. A separate MINACE filter is synthesized for each person using an automated filter-synthesis algorithm that uses a training set of illumination differences of that person and a validation set of a few faces of other persons to select the MINACE filter parameter *c*. The MINACE filter for each person is a combination of training images of only that person; no false-class training is done. Different formulations of the MINACE filter and the use of two different correlation plane metrics: correlation peak value and peak-to-correlation plane energy ratio (PCER), are examined. Performance results for face verification and identification are presented using images from the CMU Pose, Illumination, and Expression (PIE) database. All training and test set images are registered to remove tilt bias and scale variations. To evaluate the face verification and identification systems, a set of impostor images (non-database faces) is used to obtain false alarm scores (P_{FA}).

Keywords: Distortion-invariant filter (DIF), face identification, face recognition, face verification, minimum noise and correlation energy (MINACE) filter

1. INTRODUCTION

Face recognition is the automatic identification of a person based on one or more images of the person's face. A face recognition system performs either verification or identification. In face verification, the system must accept or reject the identity claim made by a user based on a comparison of the user's face to only the reference for the claimed face. If the match-score (similarity measure) is above some threshold, the user's identity is verified. In face identification, the references for all faces in the database are examined and the one with the best match-score denotes the class of the input. To enable rejection of impostor faces (faces of persons not in the database and not used in training or validation), the best match-score is required to be above some threshold, otherwise the input is rejected. In face verification, if the reference for the claimed face produces an output below a threshold, the input is rejected. The rejection of unseen impostor faces which may be similar to the faces in the database is a challenging problem in face recognition.

In prior work[1], we presented a face verification system that functioned in the presence of illumination variations. It was based on the minimum noise and correlation energy (MINACE) distortion-invariant filter (DIF)[2]. A separate MINACE filter was synthesized for each person using an automated filter-synthesis algorithm that used a training set of images of that person and a set of false-class validation images (a few faces of other persons) to select the MINACE filter parameter *c*. In this work, we use different formulations of the MINACE filter and also examine the use of the peak-to-correlation plane energy ratio (PCER) as the correlation match-score metric. The automated filter-synthesis algorithm is modified as we describe in Sect. 4.2. We use fewer training set images for each filter (four vs. six) than in our earlier work[1]. Performance results are presented for both face verification and face identification using images from the CMU Pose, Illumination, and Expression (PIE) database[3].

We now discuss prior work in face recognition. Our concern in this paper is face recognition in the presence of illumination variations (specifically those in the PIE database). We have interest only in face recognition methods that can handle both illumination and pose variations. Prior DIF work using the minimum average correlation energy (MACE) filter[4] considered face verification[5] (and addressed rejection but only addressed expression differences), and face identification[6] using the illumination differences in the PIE database (*but did not address rejection*). Both works used the peak-to-sidelobe-ratio (PSR) as the correlation plane metric rather than the correlation peak value. Fisherfaces[7] have been noted to be of use for illumination variations, but only frontal pose views were considered. Eigenfaces[8] have been shown to handle pose variations (using view-based methods) but they do not perform well on illumination variations[7,9]. Many graphic techniques[10,11,12] have been applied to the PIE and other databases with pose and illumination differences. The light-fields method[10] is computationally expensive on-line, and it assumes that a precise pose estimate is available. The morphable 3D face model[11] technique requires 3-D face images from a laser scanner. The last two methods[10,11] have been applied to the PIE database. An illumination cone[12] method has also been used but it requires much storage. The fusion[13] of LDA (linear discriminant analysis) and PCA (principal component analysis) has been applied to face recognition with expression and slight pose differences present (not with illumination differences). Quotient images[14] are illumination-invariant, but the concept cannot be extended to other distortions such as pose variations. In all of this prior work that considered face identification with one exception[9], rejection of impostor faces (non-database faces) was not considered; all inputs were assigned a class label and no rejection threshold was used.

In evaluating the performance of our face verification system, if the correct filter output for a database input is ≥ a threshold *Th*, it contributes to the number of correct scores (P_C). For any impostor test face, each filter output ≥ *Th* contributes to the false-alarm rate (P_{FA}). For face identification results, the highest filter output is retained for each test input. If the highest filter produces the correct output ≥ *Th* for a face from the database, it contributes to P_C (the number of test inputs correctly classified). If the highest score for an impostor face is ≥ *Th*, it contributes to P_{FA}. In face verification, the total number of possible false alarms is much more than in the case of face identification, since an input test image can be a false alarm for any filter in face verification. This scoring of false alarms in verification is not typically used in prior work, but it should be, since an imposter can claim to be anyone (i.e., all filters should be checked for each imposter claim). When comparing our performance to that of other classifiers that are not DIFs, to obtain fair comparison results, it may be necessary to consider an imposter face as a false alarm if any of the filters gives a peak above threshold. We use the same MINACE filters and impostor images for evaluating the performance of the face verification and identification systems. None of the impostor faces are present in the training or validation stages. The rest of this paper is organized as follows. In Sect. 2, we present two different MINACE filter formulations that we consider in this work. Section 3 describes the PIE database. Section 4 describes the automated MINACE filter-synthesis algorithm that is used to synthesize the filter for each person. Performance results (P_C and P_{FA}) for the face verification and identification systems are presented in Sect. 5.

2. DIFFERENT MINACE FILTER FOMULATIONS

This section describes the version of the original MINACE filter[2] that we use. Vectors (matrices) are denoted as lower (upper) case bold letters. All data are in the Fourier Transform (FT) domain. The 2-D FT of the filter is lexicographically ordered into a column vector **h**. The 2-D FT of each training set image included in the filter is lexicographically ordered into a column vector $\mathbf{x}_i$ of the data matrix **X**. The filter is required to give a specified correlation peak value for each training set image included in the filter; these values (usually one) are specified by the elements of a column vector **u**. These *peak constraints* are described by

$$\mathbf{X}^{\mathrm{H}}\mathbf{h} = \mathbf{u} = \begin{bmatrix}1 & 1\ldots1\end{bmatrix}^{\mathrm{T}}, \tag{1}$$

where $(\ )^{\mathrm{H}}$ denotes the complex conjugate (Hermitian) transpose. To improve performance, the filter **h** is also required to minimize a combination of correlation plane energy due to training images and correlation plane energy due to distorted versions of the objects to be recognized. We use zero-mean white Gaussian noise to model the expected distortion power spectrum. We choose the energy function to be minimized as

$$E = \mathbf{h}^{\mathrm{H}}\mathbf{T}\mathbf{h}, \tag{2}$$

where **T** is a diagonal matrix whose diagonal entries are the spectral envelope of the training images and noise at each frequency. That is,

$$\mathbf{T}(k,k) = \max[\mathbf{S}(k,k), c\mathbf{N}(k,k)], \tag{3}$$

$$\mathbf{S}(k,k) = \max[\mathbf{S}_1(k,k), \mathbf{S}_2(k,k), \ldots, \mathbf{S}_{N_T}(k,k)], \tag{4}$$

where $\mathbf{S}_i$ is a diagonal matrix whose diagonal entries are the elements of the lexicographically ordered 2-D power spectrum ($|\mathrm{FT}|^2$) of training image i , N_T is the total number of training images, **N** is the identity matrix, the maximum value in **S** is normalized to one, and c ($0 \le c \le 1$) controls the variance of the noise. The Lagrange multiplier solution that minimizes the expression in Eq. (2) subject to the constraints in Eq. (1) is

$$\mathbf{h} = \mathbf{T}^{-1}\mathbf{X}(\mathbf{X}^{\mathrm{H}}\mathbf{T}^{-1}\mathbf{X})\mathbf{u}. \tag{5}$$

The filter parameter c and the images to be included in the filter are selected using a set of training images and a set of validation images (Sect. 4.2). **S** is based on all the images in the training set, not just the ones included in the filter. Thus, **S** and **T** (for a fixed c) do not change as new training images are included in the filter. This is computationally attractive as **S** for a filter needs to be computed only once while **T** has to be computed once every time c is changed. The MINACE filter can be shown to be a linear combination of training images that have been preprocessed by $\mathbf{T}^{-1/2}$. A lower value of c makes the filter minimize correlation plane energy due to training images more and makes $\mathbf{T}^{-1/2}$ emphasize higher spatial frequencies. This makes the filter more discriminative to false objects but it makes recognition of distorted versions of an object more difficult. A higher value of c makes the filter emphasize correlation plane energy due to noise (that models distortions) more and makes $\mathbf{T}^{-1/2}$ emphasize lower spatial frequencies. This improves the distortion-tolerance performance of the filter but it also makes the rejection ability of the filter worse. Thus, c trades-off recognition versus rejection performance. Since **T** in Eq. (3) is based on the *max*( ) operator, as c is varied, **T** does not change for all spatial frequencies, e.g. as c is increased, the value of **T** at higher spatial frequencies does not change.

We also consider minimizing a different energy function in Eq. (2). Instead of defining **T** to be the spectral envelope as in Eq. (3), we define **T** as[15]

$$\mathbf{T}(k,k) = (1-c)\mathbf{S}_{avg}(k,k) + c\mathbf{N}(k,k), \tag{6}$$

$$\mathbf{S}_{avg}(k,k) = \text{mean}[\mathbf{S}_1(k,k), \mathbf{S}_2(k,k), \ldots, \mathbf{S}_{N_T}(k,k)], \tag{7}$$

where the maximum value in $\mathbf{S}_{avg}$ is normalized to one. The filter solution is Eq. (5) with the new definition of **T** in Eq. (6). We refer to this as the *additive spectrum formulation*. For $c = 1$, both formulations are identical, while for $c = 0$, the *spectral envelope formulation* reduces to the minimum correlation energy (MICE) filter[2] and the average spectrum formulation reduced to the MACE filter[4]. As c ($0 \le c \le 1$) is changed, the value of **T** changes for all spatial frequencies.

3. PIE DATABASE

The subset of the PIE database[3] we used consists of images of 65 subjects taken under 21 different illumination conditions (with the room lights off). Each subject is facing the central (frontal) camera with a neutral expression. The 21 illumination differences for each subject were captured in approximately 0.7 seconds[3], thus each subject's pose and expression is almost constant across the 21 illumination differences. We converted the original 640x486 pixel color images to grayscale, since we only use the intensity information. As was noted in prior work[1], many subjects in the PIE database have their heads titled (in all images). A face recognition system constructed from these images will use the tilt of the head in addition to facial features for performing recognition and discrimination and this will improve performance. Thus, the use of titled images is not fair. And, in a real setting, different face images of the same person will have different head tilts and may be at different scales. Thus we register (align) the faces in the images to remove tilt bias and scale variations. This registration is done by manually locating the coordinates of three facial landmarks (the eyes and the center of the mouth) in each image. Automated methods to locate the eyes[16] and other facial features[17] exist. Each 640x486 pixel image was affine transformed so that the coordinates of the three facial landmarks were the same for all images. The images were then reduced to 64x64 pixels to retain the part of face containing the eyes, the nose, and the mouth. These are referred to as *registered images*. All training and test images were registered this way. *Registration makes images of different persons look more similar and rejection of impostor faces is more difficult since impostor faces are also registered*. Figure 1 shows the registered images of subject 3 for all 21 illumination differences. The illumination differences are labeled in the order in which the flashes were indexed in the PIE database. The camera flash is on the left side of the face for illumination 1 (extreme right shadow) and the right side of the face for illumination 16 (extreme left shadow). Illumination 7 represents frontal illumination.

3.1. Training and Validation Set Databases

A separate MINACE filter is synthesized for each subject. We synthesize filters for subjects 1-40. The training set for the filter for a subject consists of four illuminations (1, 7, 16, and 19) of that subject. Illumination 19 is also frontal illumination but the vertical coordinate of the camera flash is higher than that for illumination 7. Not all training set images are included in the filter (Sect. 4.2). MINACE filter synthesis uses a validation set of false-class images to select the MINACE parameter c. We use illumination 7 of three subjects 63-65 as the false class validation set for each of the 40 filters. None of the validation set images are present in the filters. No images of subject 63-65 are used in the test stage. And, illuminations 1, 7, 16, and 19 (used in the training set) are not used in the test set (even for the impostor faces). Thus, none of the training or validation set images or illuminations are present in the test set. All training and test set images are *normalized to have unit energy* before synthesizing the filters and before performing recognition correlations (Sect. 4.2).

Figure 1. Registered images of subject 3 for all 21 different illumination conditions.

4. MINACE FILTER SYNTHESIS

4.1. Match-score Metrics and Filter Size

In prior work[1], we used the peak value of the output correlation as a match-score in the test stage. The MINACE filter is designed to produce correlation peak values of one for all images included in the filter while minimizing the correlation plane energy. Thus, the filter maximizes the ratio of the correlation peak to the correlation plane energy for each image included in the filter. We therefore consider another correlation match-score metric in this work. We use the peak-to-correlation plane energy ratio (PCER) metric which is

$$\mathrm{PCER} = \frac{\text{correlation peak value}}{\sqrt{\text{average correlation plane energy}}}.$$

The correlation peak is the largest value in the central 11x11 pixel region of the correlation plane. The average energy in the full correlation plane is computed by taking the average value of the square magnitude of each pixel (including the correlation peak). The PCER should be high for images of the filter's subject and low for images of other subjects. We evaluate PCER at the largest point in the correlation plane. The MINACE filters synthesized using the correlation peak value/PCER in the training stage are evaluated using the correlation peak value/PCER in the test stage. We use circular correlations instead of linear correlations since they are faster to compute. Thus, the size of the filters is the same as the size of the training images, 64x64 pixels.

4.2. Automated MINACE Filter Synthesis

The goal of the automated MINACE filter synthesis algorithm is to use the training and validation sets to select the filter parameter c to achieve both good recognition and imposter rejection performance. The filter-synthesis algorithm iteratively adds training images to the filter until the filter recognizes all training set images with correlation peaks $\geq$ some minimum value. Thus, the selection of the images to be included in the filter is automated. To ensure good recognition performance, we do not allow all training images to be included in the filter. We evaluate the impostor rejection performance of the filter on the validation set to determine if the value of c needs to be adjusted to achieve better impostor rejection performance. The automated MINACE filter-synthesis algorithm in Ref. 1 using the correlation peak value is shown below. The MINACE filter-synthesis algorithm parameters used were chosen from preliminary data tests.

1. Initialize c to a default value of 0.0006. Set the minimum peak constraint for the training set images, *min_true_peak* and the maximum peak constraint for the false-class validation images, *max_false_peak*.
2. Create a filter using one of the images from the training set with this value of c. We use illumination difference 1 as the first image to be included in the filter.
3. Correlate this filter with the remaining images in the training set. If any of the correlation peak values are below *min_true_peak*, then add the image with the lowest correlation peak value to the filter and synthesize a new filter with the new and old image. Continuing this process ensures that the designed filter will recognize all training set images with correlation peaks ≥ some minimum (*min_true_peak*). The filter (by definition) gives correlation peak values of 1.0 with all training set images included in the filter to satisfy Eq. (1).
4. Repeat the correlation step in (3) and continue including more images from the training set as in step (3) until the filter gives correlation peaks ≥ *min_true_peak* for all images in the training set.
5. If all images in the training set are needed to create the filter, this suggests that the filter does not have good recognition capabilities, i.e., it is not likely to give high correlation peaks for test inputs. In this case, increase the value of c by $5x10^{-5}$ and repeat steps (2) – (4) with the new value of c. Increasing c will make it easier for the filter to recognize different illumination variations.
6. Correlate this filter with the images in the false class validation set. If any of the correlation peak values for the false class validation set images are too large (above *max_false_peak)*, then the filter is not discriminative enough. Decrease the value of c by $3x10^{-5}$ and repeat steps (2) – (5). We decreased c by $3x10^{-5}$ and not by $5x10^{-5}$ to prevent the filter-synthesis algorithm from getting stuck in a loop.

Once the filter-synthesis algorithm exits step (6), a filter with a good choice of c has been created which should have both good recognition and discrimination capabilities. This concludes the synthesis process for the filter. Steps (2)-(5) ensure good recognition and step (6) ensures good discrimination. When the correlation peak value is used as the match-score, we set *min_true_peak* to 0.75 and *max_false_peak* to 0.65. When PCER is used as the match-score, we set *min_true_pcer* to 13.1 and *max_false_pcer* to 10.2. These values were chosen after performing some initial tests. When the PCER metric is used, c is decreased in step (5) and increased in step (6). Increasing c makes the MINACE filter emphasize lower spatial frequencies and results in broader correlation plane outputs for true and false class test images (this increases correlation plane energy), while decreasing c has the opposite effect (the correlation plane outputs are sharper). Increasing c seems to increase the correlation plane energy for the false-class images more than for the true-class images, thus decreasing PCER for the false-class images more. Thus, increasing c should improve impostor rejection performance. Decreasing c leads to sharper correlation plane outputs thus increasing PCER for true-class images. In future work, we will present data to explain these trends in more detail.

4.3. Filter Data

MINACE filters were synthesized for subjects 1-40 using the automated MINACE algorithm (Sect. 4.2) using both **T** measures, spectral envelope and additive spectrum in Eqs. (3) and (6), and using both the correlation peak value and PCER as performance measures. Thus, 160 (40 subjects x 2 **T** measures x 2 performance measures) filters were synthesized. Figure 2 shows the spatial templates of the MINACE filters for some of the subjects (the filters were synthesized using the spectral envelope **T** definition in Eq. (3) and the correlation peak value as the performance metric). As can be seen in Fig. 2, the filter performs edge enhancement and emphasizes key facial structures such as the eyes, the nose, and the mouth. This is all done automatically. *This is why a MINACE filter can recognize faces independent of illumination variations*.

All filters required three illumination images (the maximum allowed) to be included in the filter during synthesis. In the filter-synthesis step, the extreme right shadow (illumination 1) was the first image used in all cases (by definition). In all but three cases out of 160, the next two illumination images chosen were the extreme left shadow (illumination 16) and the frontal illumination (illumination 7) or illumination 7 followed by illumination 16. In all cases, illumination differences 1 and 16 were used in filter synthesis. This is somewhat expected, but automation is needed, as illumination 16 was chosen as the third image in ten cases. In three of the cases – filter 33 synthesized using both definitions of **T** with PCER as the performance metric and filter 11 synthesized using the additive spectrum **T** definition in Eq. (6) and with PCER as the performance metric – illumination 19 was the third image included in the filter.

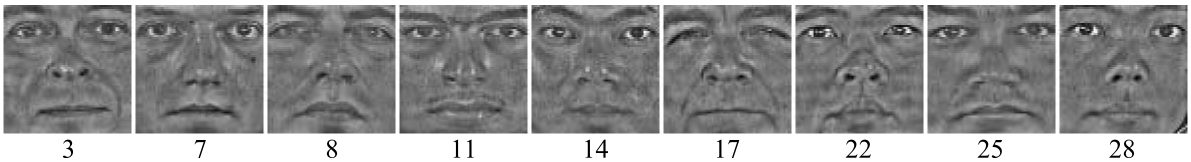

Figure 2. Spatial templates of the MINACE filters for a few subjects – key facial structures such as the eyes, the nose, and the mouth are emphasized by the filter's automatic edge enhancement.

Figure 3 shows the values of *c* used to synthesize the MINACE filters for all 40 subjects using the **T** definition in Eq. (3) and using the correlation peak value as the performance metric. The horizontal axis in Fig. 3 is the filter number and ranges from 1-40. From Fig. 2, we note that 27 (out of 40) filters were synthesized using the starting value of *c* (0.0006) and no iterations were required. The rest (13) of the filters required iterations and lower value of *c* and of those 13, the filters for subjects 23, 39, and 35 required the three lowest values of *c* ($6x10^{-5}$, $1.2x10^{-4}$, and $1.5x10^{-4}$ respectively). These subjects look more similar to one of the three validation set faces (subjects 63-65) than the other subjects do. Thus, the value of *c* for filters 23, 39, and 35 had to be decreased the most (this emphasizes higher spatial frequencies) during filter synthesis to achieve better discrimination against the impostor faces in the validation set. When PCER was used as the performance metric, most filters required a different value of *c* (≈ 0.0003), half of the prior value. A different value of *c* is expected because the performance metric used during filter synthesis was different. We also note that for each subject and for each correlation plane metric (peak or PCER) considered, the *c* value used to synthesize the MINACE filter using the two definitions of **T** in Eqs. (3) and (6) were either the same or differed by small amounts. For example, when the correlation peak value was used as the performance measure, for 29 (out of 40) subjects, the filter synthesized using the spectral envelope definition of **T** in Eq. (3) required the same value of *c* as the filter synthesized using the additive spectrum definition of **T** in Eq. (6). The maximum difference in *c* value was $2.1x10^{-4}$ (subject 28).

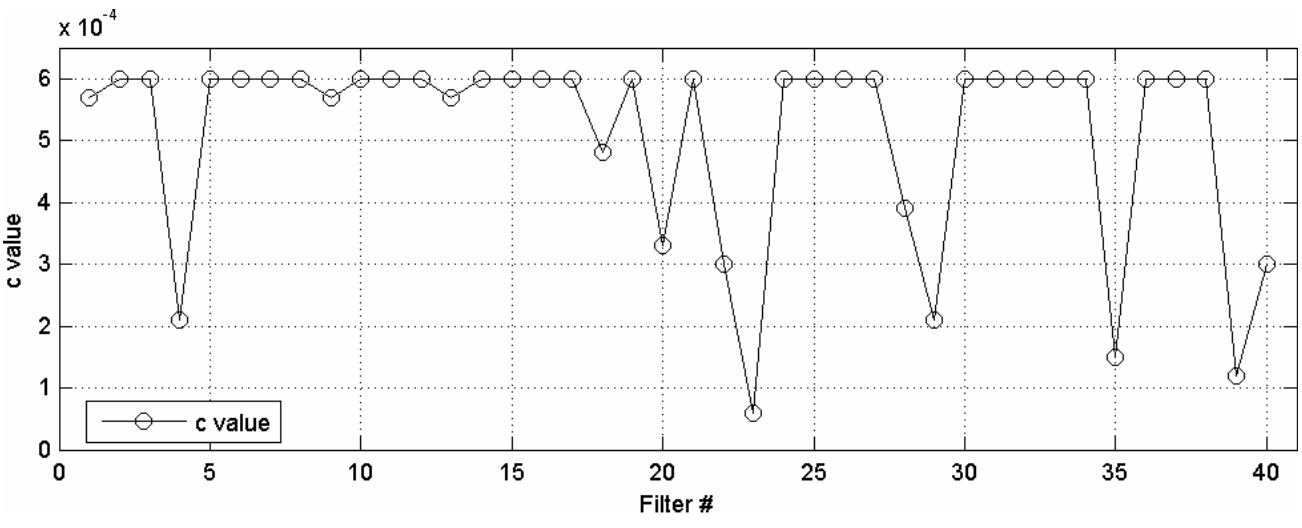

Figure 3. *c* values used to synthesize the filters for subjects 1-40 using the spectral envelope definition of **T** in Eq. (3) and using the correlation peak value as the performance metric; different filters require different values of *c* during filter synthesis.

5. PERFORMANCE RESULTS

5.1. Test Set

For face verification, the true-class test set for the filter for each of the 40 subjects consists of 17 true-class images (the 17 illuminations out of 21 that were not used in the training set) for that subject. As the false-class test set for each filter, we use face images of the 22 subjects 41-62 each with the 17 illumination differences that were not seen in the training stage. The MINACE filter for each subject should reject these impostor images. When considering face identification, the test set images to be classified are the 17 illumination differences of subjects 1-40 (that were not seen in the training stage). For both face verification and identification results, P_C is a percentage out of 680 (17x40) test set images. For face verification, P_{FA} is a percentage out of 14960 images (17 illumination differences x 22 impostor faces x 40 filters), while for face identification it is a percentage out of 374 (17x22) images.

5.2. Face Verification Results

Figure 4 shows the minimum true (authentic) test set peak (curve ○) and the maximum false (impostor) test set peak (curve x) for the filters for each of subjects 1-40 synthesized using the spectral envelope definition of **T** in Eq. (3) and using the correlation peak value as the performance measure. The horizontal axis in Fig. 4 is the filter number and ranges from 1-40. In Fig. 4, we also show the overall minimum authentic test set peak (indicated by '–') and the overall maximum impostor test set peak (indicated by '– ·'). We wish to use the same correlation peak threshold for all filters in the test stage to evaluate performance scores (this is typically required). From Fig. 4, we see that the overall minimum authentic test set peak (0.594 for filter 23) is below the overall maximum impostor test set peak (0.856 for filter 7), thus the use of any fixed correlation peak threshold in the test stage will result in either authentic faces being rejected (misses) or impostor faces being recognized (false alarms) or both. For example, if we use a correlation peak threshold of 0.86 in the test stage, we can achieve P_{FA} = 0%, however, the filters for 12 subjects (1, 2, 4, 12, 14, 20, 23, 29, 35-37, and 39) will reject authentic faces using this threshold (the P_C score is 91.76%, since most illuminations of each subject are above threshold). We note that if we use different thresholds for some of the filters, e.g. if we use a threshold = 0.59 for filter 23, we can improve performance results, increasing the P_C for filter 23 from 47.06% (8 misses out of 17) to 100%. With a different threshold allowed for each filter, we can obtain P_C = 100% and P_{FA} = 0%, but use of different thresholds is typically not allowed. We also note that for the filters for subjects 1-40 synthesized using the spectral envelope definition of **T** in Eq. (3) and using PCER as the performance metric, the overall minimum authentic test set PCER (9.82) is greater than the overall maximum impostor test set PCER (9.75). Thus for a range of PCER test thresholds from 9.76-9.82 (the same threshold is applied to all filters), we achieve P_C = 100% and P_{FA} = 0%. *Thus, PCER is a more preferable metric than correlation peak value*.

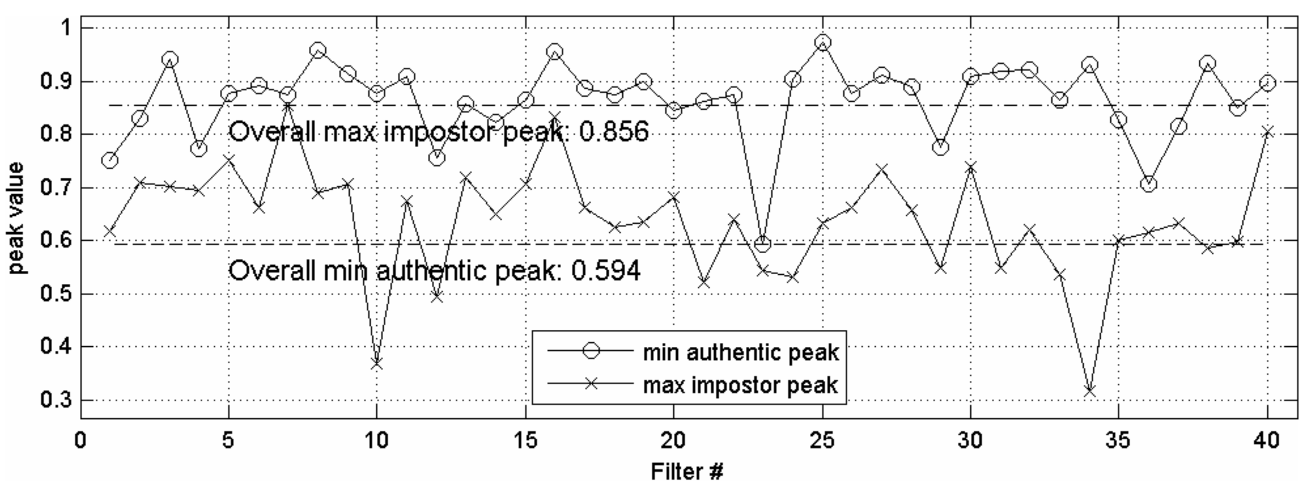

Figure 4. Minimum authentic and maximum impostor test set correlation peak values for each of the MINACE filters for subjects 1-40 synthesized using the spectral envelope definition of **T** in Eq. (3) and using the correlation peak value as the performance metric.

To compare the performance of the face verification systems using MINACE filters with different definitions of **T** and using different correlation plane metrics, we plot ROC curves of P_C vs. P_{FA} for different values of the correlation match-score threshold. Figure 5a shows the ROC curves of P_C vs. P_{FA} for the MINACE filters synthesized using the spectral envelope definition of **T** in Eq. (3) and using both the correlation peak value (the lower curve ○) and PCER (the upper curve Δ) as the performance metrics. The horizontal axis in Fig. 5a is P_{FA} and the vertical axis is P_C. A value of 0.0067 along the horizontal axis indicates a total of one false alarm for all 14960 total imposter inputs for all 40 filters, while a value of 0.147 along the vertical axis indicates a total of one correct verification out of a total of 680 true inputs for all 40 filters. In Fig. 5, we also show our version of the equal error rate (EER) point (*) for the two curves; in this case, it is where the total number of misses (total number of true-class test set images rejected) equals the total number of false alarms, along with the match-score threshold used in the test stage. Note that we compute the EER point by considering the total number of errors (misses or false alarms) rather than the more conventional percentages of errors. From Fig. 5a, we see that the correlation peak ○ curve lies below and to the right of the PCER Δ curve. That is, for a given P_{FA} score, the P_C score for the PCER metric is higher, and for a given P_C score, the P_{FA} score for the PCER metric is lower. For example for P_{FA} = 0.2%, the P_C scores for the correlation peak and PCER curves are 98.24% and 100% respectively, and for P_C = 99%, the P_{FA} scores are 0.33% and 0% respectively. Thus, *the performance scores for the MINACE filters synthesized and tested using PCER as the performance metric are clearly better than the scores for the MINACE filters synthesized and tested using the correlation peak value as the performance metric*.

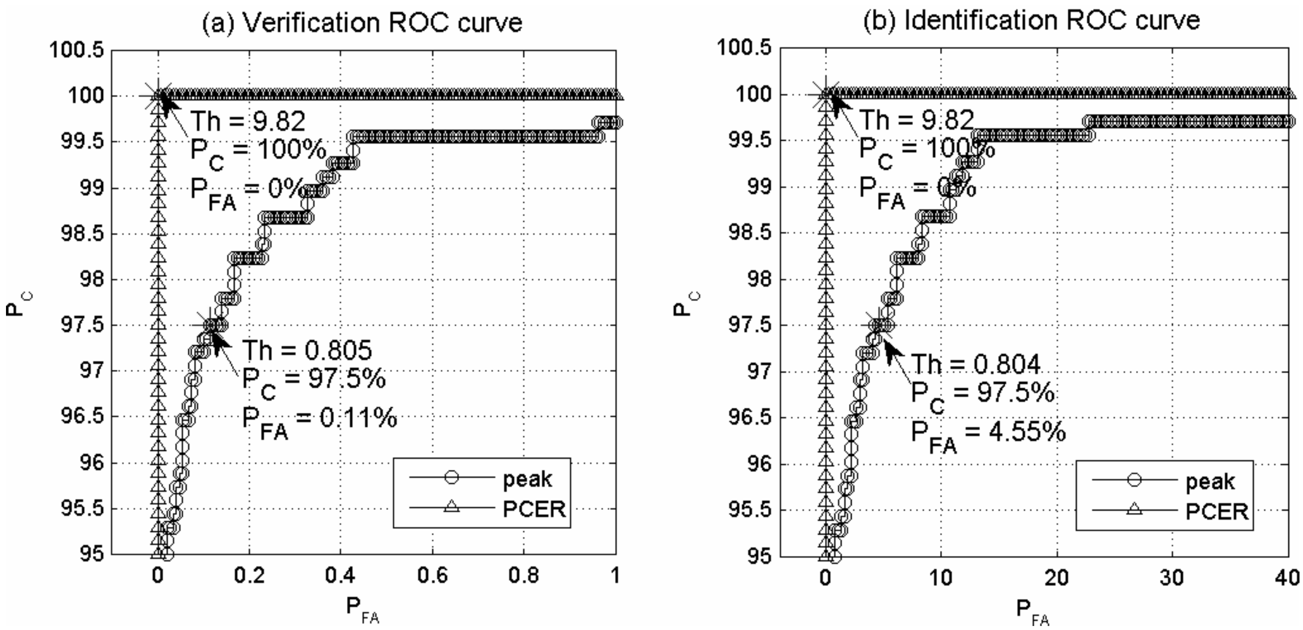

Figure 5. ROC curves (P_C vs. P_{FA}) for the MINACE filters for subjects 1-40 synthesized using the spectral envelope definition of **T** in Eq. (3) and using both the correlation peak value (lower curve ○) and PCER (upper curve Δ) as performance metrics for face (a) verification and (b) identification. The EER points where the total number of misses equals the total number of false alarms are indicated by * and the corresponding thresholds and P_C and P_{FA} scores are shown.

We now compare the P_C and P_{FA} scores at the EER point for the MINACE filters for subjects 1-40 synthesized using both the spectral envelope and the additive spectrum definitions of **T** in Eqs. (3) and (6) and using the correlation peak value and PCER as the correlation plane metrics. Table 1 shows these values along with the match-score threshold required in the test stage to achieve the EER operating point. From Table 1, we see that for a given **T** definition (spectral envelope or additive spectrum), the performance of the systems using the PCER as the performance metric is much better than the performance of the system using the correlation peak value as the performance metric, e.g. for the spectral envelope formulation, 0 total misses using the PCER versus 34 total misses using the correlation peak. We also note that using the PCER metric, both MINACE filter **T** formulations produce very similar results (0 total errors for the spectral envelope formulation versus two total errors for the additive spectrum formulation). However, when using the correlation peak value as the performance metric, the use of the additive spectrum formulation rather than the spectral envelope formulation leads to a reduction of six misses and six false alarms (a total of 12 errors) at the EER point. This suggests that *one should use the additive spectrum definition of* **T** when using the correlation peak value as the performance metric. However, we will have to examine this in more detail in future work.

Table 1. Comparison face *verification* P_C and P_{FA} scores for the MINACE filters for subjects 1-40 synthesized using both the spectral envelope and additive spectrum definition of **T** in Eqs. (3) and (6) and using the correlation peak value and PCER as performance measures at the EER point where the total number of misses equals the total number of false alarms and the test threshold used.

Performance metric	Peak		PCER	
MINACE formulation	Spectral envelope	Additive spectrum	Spectral envelope	Additive spectrum
Test threshold	0.805	0.78	9.82	10.02
P_C	97.5% 17 misses	98.38% 11 misses	100% 0 misses	99.85% 1 miss
P_{FA}	0.11% 17 false alarms	0.07% 11 false alarms	0% 0 false alarms	0.007% 1 false alarm

5.3. Face Identification Results

To compare the performance of the face identification systems using MINACE filters with different definitions of **T** and using different correlation plane metrics, we plot ROC curves of P_C vs. P_{FA} for different values of the correlation match-score threshold. Figure 5b shows the ROC curves of P_C vs. P_{FA} for the MINACE filters synthesized using the spectral envelope definition of **T** in Eq. (3) and using both the correlation peak value (lower curve ○) and PCER (upper curve Δ) as the performance metrics. The horizontal axis in Fig. 5b is P_{FA} and the vertical axis is P_C. A value of 0.267 along the horizontal axis indicates a total of one false alarm out of a total of 374 (17x22) test set impostor faces and a value of 0.147 along the vertical axis indicates a total of one correct classification out of a total of 680 (17x40) true test set database faces. In Fig. 5b, we also show our equal error rate (EER) point. As was the case in face verification, the correlation peak ○ curve lies below and to the right of the PCER Δ curve and *the MINACE filters synthesized and tested using the PCER as the performance metric perform much better than the MINACE filters synthesized and tested using the correlation peak value as the performance metric.*

We now compare the P_C and P_{FA} scores at the EER point for the MINACE filters for subjects 1-40 synthesized using both the spectral envelope and additive spectrum definitions of **T** in Eqs. (3) and (6) and using the correlation peak value and PCER as the correlation plane metrics. Table 2 shows these values along with the match-score threshold required in the test stage to achieve the EER operating point. In Table 2, we also show the maximum classification score, P_{CM}, when no rejection threshold is used. A comparison of Tables 1 and 2 shows that the total number of misses and the total number of false alarms for verification and identification are identical at the EER point for both MINACE filter formulations and both match-score metrics, e.g., 11 misses and 11 false alarms using the additive spectrum formulation and using the correlation peak as the performance metric for both face verification and identification. This is expected since P_C is the same in both tests and it sets the number of errors. The P_{CM} scores using the correlation peak value as the performance measure are 99.71% (for both MINACE formulations), that is, two misclassifications out of 680 (17x40) test inputs, while the P_{CM} scores using the PCER as the performance metric are 100% for both MINACE formulations. Thus, if rejection of impostor faces is not of concern, then any of the combinations of the two MINACE filter formulations and the two correlation match-scores can be used to build the face verification system since all of them achieve 100% or nearly 100% accuracy.

Table 2. Comparison face *identification* P_C and P_{FA} scores for the MINACE filters for subjects 1-40 synthesized using both the spectral envelope and additive spectrum definition of **T** in Eqs. (3) and (6) and using the correlation peak value and PCER as performance measures at the EER point where the total number of misses equals the total number of false alarms and the test threshold used. Also shown is P_{CM}, the maximum classification score when no rejection threshold is used.

Performance metric	Peak		PCER	
MINACE formulation	Spectral envelope	Additive spectrum	Spectral envelope	Additive spectrum
Test threshold	0.804	0.778	9.82	10.02
P_C	97.5% 17 misses	98.38% 11 misses	100% 0 misses	99.85% 1 miss
P_{FA}	4.55% 17 false alarms	2.94% 11 false alarms	0% 0 false alarms	0.27% 1 false alarm
P_{CM}	99.71% 2 misclassifications	99.71% 2 misclassifications	100% 0 misclassifications	100% 0 misclassifications

6. SUMMARY

Two different MINACE filter formulations (spectral envelope and additive spectrum) and two different correlation plane metrics (peak and PCER) were used to create face recognition systems that function with illumination variations present. Performance scores were presented for both face verification and identification. Using the spectral envelope definition of the MINACE filter and using PCER as the performance measures, perfect results (P_C = 100% and P_{FA} = 0%) were obtained for both face verification and identification. This is better than our earlier work where we achieved an EER of seven misses and seven false alarms for face verification (we did not consider face identification in that work), which corresponded to P_C = 99.28% and P_{FA} = 0.01%. And, in comparison to our prior work, we used two fewer illumination differences per subject (four vs. six) in the training set. Note that the MINACE filter for each subject has not seen the other database faces (to be classified) during the training stage. This is different from other work[7] where the decision to classify or reject a test database face is made using all database faces during training. The advantage of using DIF based methods such as the MINACE filter is that when new faces (to be classified) are included in the database, the MINACE filters for existing faces do not have to be resynthesized, while other methods[7] have to recompute the templates or parameters for existing database faces.

We noted that *the spectral envelope and additive spectrum MINACE filter formulations produced very similar results for both face verification and identification when PCER was used as the performance metric. However, when the correlation peak value was used as the performance metric, the performance results were much better using the additive spectrum formulation*. We need to examine the reasons for this in future work. The use of the additive spectrum formulation and the use of the correlation peak value as the performance metric produced an EER of 11 misses and 11 false alarms for both face verification and face identification which are significantly worse than the perfect scores (0 misses and 0 false alarms) obtained using the spectral envelope formulation and using PCER as the performance metric. In future work, we will examine the use of a combination of the correlation peak value and PCER as the performance metric during MINACE filter synthesis and for evaluating the face recognition system. We will also extend our face recognition system to handle pose variations in addition to illumination variations.

REFERENCES

1. R. Patnaik and D. P. Casasent, "Face verification and rejection with illumination variations using MINACE filters," in *Optical Pattern Recognition XV*, D. P. Casasent, T. Chao, eds., *Proc. SPIE* **5437**, pp 277-287, April 2004.
2. G. Ravichandran and D. P. Casasent, "Minimum noise and correlation energy optical correlation filter," *Applied Optics* **31**(11), pp. 1823-1833, April 1992.
3. T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination and Expression (PIE) database," *Proc. Fifth IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, pp. 46-51, May 2002.

4. A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, “Minimum average correlation energy filters,” *Applied Optics* **26**(17), pp. 3633-3640, September 1987.
5. M. Savvides, B.V.K. Vijaya Kumar, and P. K. Khosla, "Face verification using correlation filters," *Proc. Third IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 56-61, March 2002.
6. M. Savvides and B. V. K. Vijaya Kumar, “Efficient design of advanced correlation filters for robust distortion-tolerant face recognition,” *Proc. IEEE Conf. on Advanced Video and Signal Based Surveillance*, pp. 45-52, July 2003.
7. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence* **19**(7), pp. 711-720, July 1997.
8. A. Pentland, B. Moghaddam, and T. Starner, “View-based and modular Eigenspaces for face recognition,” *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 84-91, June 1994.
9. D. P. Casasent and C. Yuan, “Face recognition with pose and illumination variations using new SVRDM support vector machine,” in *Optical Pattern Recognition XV*, D. P. Casasent, T. Chao, eds., *Proc. SPIE* **5437**, pp. 1-12, April 2004.
10. R. Gross, I. Matthews, and S. Baker, “Appearance-based face recognition and light-fields,” *IEEE Trans. on Pattern Analysis and Machine Intelligence* **26**(4), pp. 449-465, April 2004.
11. V. Blanz, S. Romdhani, and T. Vetter, “Face identification across different poses and illuminations with a 3D morphable model,” *Proc. Fourth IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, pp. 192-197, March 2002.
12. A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, “From few to many: Illumination cone models for face recognition under variable lighting and pose,” *IEEE Trans. on Pattern Analysis and Machine Intelligence* **23**(6), pp. 643-660, June 2001.
13. G. L. Marcialis and F. Roli, “Fusion of LDA and PCA for face verification,” *Proc. Seventh European Conf. on Computer Vision*, vol. LNCS **2359**, pp. 30–37, June 2002.
14. A. Shashua and T. Riklin-Raviv, “The quotient image: Class-based re-rendering and recognition with varying illuminations,” *IEEE Trans. on Pattern Analysis and Machine Intelligence* **23**(2), pp. 129-139, February 2001.
15. D. Casasent and S. Ashizawa, “Synthetic aperture radar detection, recognition, and clutter rejection with new minimum noise and correlation energy filters,” *Optical Engineering* **36**(10), pp. 2729-2736, October 1997.
16. K. Lam and H. Yan, “Locating and extracting the eye in human face images,” *Pattern Recognition* **29**(5), pp. 771-779, May 1996.
17. R. Chellappa, C. L. Wilson, and S. Sirohey, “Human and machine recognition of faces: A survey,” *Proc. IEEE* **83**(5), pp. 705-741, May 1995.