

# THE CORRENTROPY MACE FILTER FOR IMAGE RECOGNITION

*Kyu-Hwa Jeong, Jose C. Principe*

Computational NeuroEngineering Laboratory  
Department of Electrical and Computer Engineering  
University of Florida, Gainesville, FL 32611 U.S.A.  
E-mail: {khjeong, principe}@cnel.ufl.edu

## ABSTRACT

The minimum average correlation energy (MACE) filter is a well known correlation filter for pattern recognition. This paper proposes a nonlinear extension to the MACE filter using the recently introduced correntropy function in feature space. Correntropy is a positive definite function that generalizes the concept of correlation by utilizing higher order moment information of signal structure. Since the MACE is a spatial matched filter for an image class, the correntropy MACE can potentially improve its performance. We apply the correntropy MACE filter to face recognition and show that the proposed method indeed outperforms the traditional linear MACE in both generalization and rejection abilities.

## 1. INTRODUCTION

Correlation filters have been applied successfully to target detection and recognition problems such as automatic target recognition (ATR) [1] and face recognition [2]. Object recognition is performed by cross-correlating an input image with a synthesized template (filter) and the correlation output is searched for the peak, which is used to determine whether the object of interest is present or not. It is well known that matched filters are the optimal linear filters for signal detection under linear channel and white noise conditions [3]. For image detection, matched spatial filters (MSF) are optimal in the sense that they provide the maximum output signal to noise ratio (SNR) for the detection of a known image in the presence of white noise, under the reasonable assumption of Gaussian statistics [4]. However, the performance of the MSF is very sensitive to even small changes in the reference image and the MSF cannot be used for multiclass pattern recognition since the MSF is only optimum for a single image. Therefore distortion invariant composite filters have been proposed in various papers [1].

The most well known of such composite correlation filters are the synthetic discriminant function (SDF) [5] and its

variations. In the conventional SDF approach, the filter is matched to a composite image that is a linear combination of the training image vectors such that the cross correlation output at the origin has the same value with all training images. The hope is that this composite image will correlate equally well not only with the training images but also with other distorted versions of that training images, even with test images in the same class. The shortcomings of the conventional SDF are that the SDF does not consider any input noise and it has a poor rejecting ability for out-of-class (false) images since it controls only a single point in the output correlation plane. Minimum variance SDF (MVSDF) filter has been proposed in [6] taking into consideration additive input noise. The MVSDF minimizes the output variance due to zero-mean input noise while satisfying the same linear constraints as the SDF. One of major difficulty in MVSDF is that we often do not know the noise covariance exactly; even when we do know it, we need its inversion and it may be computationally impossible in practice. Another correlation filter that is widely used is the minimum average correlation energy (MACE) filter [7]. The MACE minimizes the average correlation energy of the output over the training images to produce a sharp correlation peak subject to the same linear constraints as the MVSDF and SDF filters. In practice, the MACE filter performs better than the MVSDF with respect to rejecting out-of-class input images. The MACE filter; however, has been shown to have poor generalization properties, that is, images in the recognition class but not in the training exemplar set are not recognized well. Therefore, the optimal trade-off filters have been proposed by [8] to combine the properties of various SDF's. Most of these are linear correlation filters. A nonlinear extension to the MACE filter has been proposed in [9].

Recently, kernel based learning algorithms have been exploited due to the fact that linear algorithms can be easily extended to nonlinear versions by the kernel method [10]. The kernel matched spatial filter (KMSF) has been proposed for hyperspectral target detection in [11] and the kernel SDF has been proposed for face recognition [12]. A new generalized correlation function, called correntropy, defined in a nonlinear reproducing kernel Hilbert space (RKHS) has been proposed

---

This work was supported in part by the National Science Foundation under grant ECS-0300340 and ECS-0601271

sed in [13] and its application to the matched filter has been presented in [14].

In this paper we propose a new nonlinear version of MACE filter based on correntropy that has better generalization and rejecting performance than the conventional MACE. In section 2, the MACE filter is reviewed briefly. In section 3, we propose a new MACE filter based on correntropy. In section 4, we present simulation results for face recognition and section 5 summarizes and points out some further research.

## 2. MINIMUM AVERAGE CORRELATION ENERGY FILTER

We consider a 2-dimensional  $i$ th image as a  $d \times 1$  column vector  $\mathbf{x}_i$ , where  $d$  is the number of pixel. This 1-dimensional discrete sequence can be obtained by lexicographically reordering the image. The conventional MACE filter is formulated in frequency domain.

The discrete Fourier transform (DFT) of the sequence  $\mathbf{x}_i$  is denoted by  $\mathbf{X}_i$  and we define the training image data matrix  $\mathbf{X}$  as

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N], \quad (1)$$

where the size of  $\mathbf{X}$  is  $d \times N$  and  $N$  is the number of training data. Let the vector  $\mathbf{h}$  be the filter in space domain and its Fourier transform vector be  $\mathbf{H}$ . We are interested in the correlation of the input image and the filter. The correlation of the  $i$ th image sequence  $x_i(n)$  with filter sequence  $h(n)$  can be written as

$$g_i(n) = x_i(n) \otimes h(n). \quad (2)$$

By Parseval's theorem, the correlation energy of the  $i$ th image can be written as a quadratic form

$$E_i = \mathbf{H}^H \mathbf{D}_i \mathbf{H}, \quad (3)$$

where  $\mathbf{D}_i$  is a diagonal matrix of size  $d \times d$  whose diagonal elements are the magnitude squared of the associated element of  $\mathbf{X}_i$ , that is, the power spectrum of  $x_i(n)$  and the superscript  $H$  denotes the Hermitian transpose. The objective of the MACE filter is to minimize the average correlation energy over all signals while simultaneously satisfying intensity constraint at the origin for each signal. The value of the correlation at the origin can be written as

$$g_i(0) = \mathbf{X}_i^H \mathbf{H} = c_i, \quad (4)$$

for all  $i = 1, 2, \dots, N$  training images, where  $c_i$  is the user specified output correlation value at the origin for the  $i$ th image. Then the average energy over all training images is expressed as

$$E_{avg} = \mathbf{H}^H \mathbf{D} \mathbf{H}, \quad (5)$$

where

$$\mathbf{D} = \frac{1}{N} \sum_{i=1}^N \mathbf{D}_i. \quad (6)$$

The MACE design problem is to minimize  $E_{avg}$  while satisfying the constraint,  $\mathbf{X}^H \mathbf{H} = \mathbf{c}$ , where  $\mathbf{c} = [c_1, c_2, \dots, c_N]$  is an  $N$  dimensional vector. This optimization problem can be solved using Lagrange multipliers, and the solution is

$$\mathbf{H} = \mathbf{D}^{-1} \mathbf{X} (\mathbf{X}^H \mathbf{D}^{-1} \mathbf{X})^{-1} \mathbf{c}. \quad (7)$$

It is clear that  $\mathbf{h}$  can be obtained from  $\mathbf{H}$  by an inverse DFT. Once  $\mathbf{h}$  is determined, we apply an appropriate threshold to the output correlation plane and decide on the class of the test image.

## 3. NONLINEAR VERSION OF THE MACE IN RKHS USING CORRENTROPY

### 3.1. Correntropy Function in RKHS

Correntropy, as proposed in [13], is a positive definite function that generalizes the correlation function to nonlinear (non Gaussian) manifolds. The correntropy of the random process  $x(n)$  at instances  $i$  and  $j$  is defined as

$$v(i, j) = E[k(x(i), x(j))], \quad (8)$$

where  $E$  is the expectation operator and  $k$  is a kernel function that obeys the Mercer's conditions. In this paper, we use the Gaussian kernel, which is the most widely used Mercer kernel,

$$k(\mathbf{x}, \mathbf{y}) = \frac{1}{\sqrt{2\pi\sigma}} \exp - \left( \frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2} \right). \quad (9)$$

For the discrete-time sequence  $\mathbf{x}$ , we can estimate the correntropy as

$$\hat{v}_{xx}(m) = \frac{1}{M - m + 1} \sum_{l=m}^M k(x_l - x_{l-m}), \quad (10)$$

where  $M$  is the number of samples of  $\mathbf{x}$ .

Correntropy has very nice properties that make it useful for nonlinear signal processing. First and foremost, it is a positive function, which means that it also defines a RKHS, but unlike the RKHS defined by the covariance function of random process it contains higher order statistical information. This new function quantifies the average angular separation in the kernel feature space of the random process at a given temporal lag. Therefore, correntropy can be the metric for similarity measurement in feature space.

According to [15], there exists a mapping  $f$  in a stationary stochastic process  $x(n)$  such that

$$v(i, j) = E[k(x(i), x(j))] = E[f(x(i))f(x(j))], \quad (11)$$

for a class of joint probability density functions (PDFs) of  $(x(i), x(j))$ , that is, there exists a nonlinear mapping  $f$  which makes the correntropy of  $x(n)$  the correlation of  $f(x(n))$ . Equation (11) allows the replacement of the correlation by the correntropy function. The proof is in [15].

### 3.2. Minimum Average Correntropy Energy Filter : Correntropy MACE

Let the  $i$ th image vector be  $\mathbf{x}_i = [x_i(1)x_i(2) \cdots x_i(d)]^T$  and filter be  $\mathbf{h} = [h(1)h(2) \cdots h(d)]^T$ , where  $T$  denotes transpose. Here, the correntropy MACE filter is formulated in the feature space by applying a nonlinear mapping function  $f$  onto the data as well as filter. We denote the transformed training image matrix and filter vector whose size are  $d \times N$  and  $d \times 1$ , respectively, be

$$\mathbf{F}_X = [\mathbf{f}_{x1}, \mathbf{f}_{x2}, \cdots, \mathbf{f}_{xN}], \quad (12)$$

$$\mathbf{f}_h = [f(h(1))f(h(2)) \cdots f(h(d))]^T. \quad (13)$$

where,

$$\mathbf{f}_{xi} = [f(x_i(1))f(x_i(2)) \cdots f(x_i(d))]^T \quad (14)$$

for  $i = 1, 2, \dots, N$ . Based on (11) we can estimate the cross correntropy between  $i$ th training image vector and the filter with given samples as

$$v_{oi}[m] = \frac{1}{d} \sum_{n=1}^d f(h(n))f(x_i(n-m)), \quad (15)$$

for all the lags  $m = -d+1, -d+2, \dots, 0, 1, \dots, d-1$ . Then we can form a cross correntropy vector  $\mathbf{v}_{oi}$  including all the lags of  $v_{oi}[m]$  denoted by

$$\mathbf{v}_{oi} = \mathbf{S}_i \mathbf{f}_h, \quad (16)$$

where,  $\mathbf{S}_i$  is the matrix of size  $(2d-1) \times d$  as

$$\mathbf{S}_i = \begin{pmatrix} f(x_i(d)) & 0 & \dots & \dots & 0 \\ f(x_i(d-1)) & f(x_i(d)) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f(x_i(1)) & f(x_i(2)) & \dots & \dots & f(x_i(d)) \\ 0 & f(x_i(1)) & \dots & \dots & f(x_i(d-1)) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & f(x_i(1)) \end{pmatrix}, \quad (17)$$

Since the scale factor  $1/d$  has no influence on the solution, we ignore the scale factor in this paper.

Then the correntropy energy of the  $i$ th image is given by

$$E_i = \mathbf{v}_{oi}^T \mathbf{v}_{oi} = \mathbf{f}_h^T \mathbf{S}_i^T \mathbf{S}_i \mathbf{f}_h. \quad (18)$$

Here, we denote  $\mathbf{S}_i^T \mathbf{S}_i = \mathbf{V}_{xi}$  and by using the definition of correntropy in (11), we can obtain the  $d \times d$  correntropy matrix  $\mathbf{V}_{xi}$  as

$$\mathbf{V}_{xi} = \begin{pmatrix} v_{xi}(0) & v_{xi}(1) & \dots & v_{xi}(d-1) \\ v_{xi}(1) & v_{xi}(0) & \dots & v_{xi}(d-2) \\ \vdots & \vdots & \ddots & \vdots \\ v_{xi}(d-1) & v_{xi}(d-2) & \dots & v_{xi}(0) \end{pmatrix}, \quad (19)$$

where, each element of the matrix is computed without explicitly knowing the mapping function  $f$  by

$$v_{xi}(l) = \sum_{n=1}^d k(x_i(n), x_i(n+l)), \quad (20)$$

for  $l = 0, 1, \dots, d-1$ .

The average correntropy energy over all the training data can be written as

$$E_{av} = \frac{1}{N} \sum_{i=1}^N E_i = \mathbf{f}_h^T \mathbf{V}_X \mathbf{f}_h, \quad (21)$$

where

$$\mathbf{V}_X = \frac{1}{N} \sum_{i=1}^N \mathbf{V}_{xi}. \quad (22)$$

Since our objective is to minimize the average correntropy energy in the linear feature space, we can formulate the optimization problem by

$$\min \mathbf{f}_h^T \mathbf{V}_X \mathbf{f}_h, \quad \text{subject to } \mathbf{F}_X^T \mathbf{f}_h = \mathbf{c}. \quad (23)$$

where,  $\mathbf{c}$  is the desired vector for all the training images. The constraint in (23) means that we specify the correntropy values between the training input and the filter as the desired constant. Since the correntropy matrix  $\mathbf{V}_X$  is positive definite, there exists an analytic solution. Then the solution in feature space becomes

$$\mathbf{f}_h = \mathbf{V}_X^{-1} \mathbf{F}_X (\mathbf{F}_X^T \mathbf{V}_X^{-1} \mathbf{F}_X)^{-1} \mathbf{c}. \quad (24)$$

In order to test this filter, let  $\mathbf{Z}$  be the matrix of  $L$  vector testing images, then the  $L \times 1$  output vector is given by

$$\mathbf{y} = \mathbf{F}_Z^T \mathbf{V}_X^{-1} \mathbf{F}_X (\mathbf{F}_X^T \mathbf{V}_X^{-1} \mathbf{F}_X)^{-1} \mathbf{c}. \quad (25)$$

Since we do not explicitly know the nonlinear mapping function  $f$ , the final output expression is obtained by approximating  $f(x(i))f(x(j))$  by  $k(x(i), x(j))$ , which holds good on an average because of (11). Hence, we do not need to find the transformation  $f(\cdot)$  as expected by the "kernel trick". In practice, the drawback of the proposed correntropy MACE filter is on its computation complexity. Fortunately,  $\mathbf{V}_X$  is a Toeplitz matrix and  $\mathbf{F}_X^T \mathbf{V}_X^{-1} \mathbf{F}_X$  is a symmetric matrix, therefore these special structures may be used to reduce the computational complexity of its inversion. More mathematical and numerical concerns are needed to obtain a fast algorithm for implementation of the correntropy MACE filter.

Applying an appropriate threshold to the output of (25), one can detect and recognize the testing data without generating the composite filter in the feature space.

### 3.3. Prewhitening in Feature Space

The MACE filter can be decomposed as a cascade of the preprocessor and the projection SDF. The preprocessor forces

the average power spectrum of the training images to become white. The proposed nonlinear version of the MACE also has the same property in feature space.

We denote  $\mathbf{F}_X^- = \mathbf{V}_X^{-1/2} \mathbf{F}_X$ , and due to the property of correntropy matrix (Toeplitz), we can decompose  $\mathbf{f}_h$  to

$$\begin{aligned} \mathbf{f}_h &= \mathbf{V}_X^{-1/2} \mathbf{V}_X^{-1/2} \mathbf{F}_X (\mathbf{F}_X^T \mathbf{V}_X^{-1/2} \mathbf{V}_X^{-1/2} \mathbf{F}_X)^{-1} \mathbf{c} \\ &= \mathbf{V}_X^{-1/2} \mathbf{F}_X^- (\mathbf{F}_X^-^T \mathbf{F}_X^-)^{-1} \mathbf{c} \end{aligned} \quad (26)$$

Equation (26) implies that the training data can be whitened in feature space by the correntropy matrix. In the space domain, the autocorrelation matrix can be used as a preprocessor for prewhitening. In feature space, intuitively, we can also expect that the correntropy matrix can be used for prewhitening. However, in practice, we cannot obtain whitened data explicitly since we do not know the mapping function.

#### 4. SIMULATIONS

In this section, we show the performance of the proposed correntropy MACE filter for face image recognition. In the simulations, we used the facial expression database collected at the Advanced Multimedia Processing Lab at the Electrical and Computer Engineering Department of Carnegie Mellon university [16]. The database consists of 13 subjects, whose facial images were captured with 75 varying expressions. The size of each image is  $64 \times 64$ . Sample images are depicted in Fig. 1. In this paper, we tested the proposed method with the original database images as well as with noisy images. Sample images with additive Gaussian noise with a 10 dB SNR are shown in Fig. 1(c). We used only 5 images to composite template (filter) per person ( the MACE filter shows a reasonable recognition result with a small number of training image in this database [2]). In order to evaluate the performance of the MACE filter in this data set, we examined 975( $13 \times 75$ ) correlation outputs. From these results and the ones reported in [2] we picked and report the results of the two most difficult cases who produced the worst performance with the conventional MACE method. We test with all the images of each person's data set resulting in 75 outputs for each class. The simulation results have been obtained by averaging (Monte-Carlo approach) over 100 different training sets (each training set consists of randomly chosen 5 images) to minimize the problem of performance differences due to splitting the relatively small database in training and testing sets. The kernel size,  $\sigma$ , is chosen to be 10 for the correntropy matrix during training and 30 for test output. In this data set, it has been observed that the kernel size around 30%-50% of the standard deviation of the input data would be appropriate. Moreover, we can control the performance by choosing a different kernel size during training for prewhitening.

Fig. 2 shows the average test output peak values for image recognition. The desired output peak value should be close to one when the test image belongs to the training image class

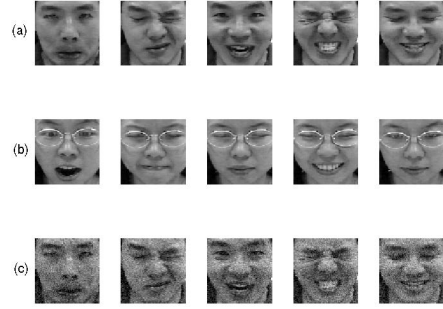


Fig. 1. Sample images: (a) True class images (b) False class images (c) True class images with additive Gaussian noise (SNR=10db).

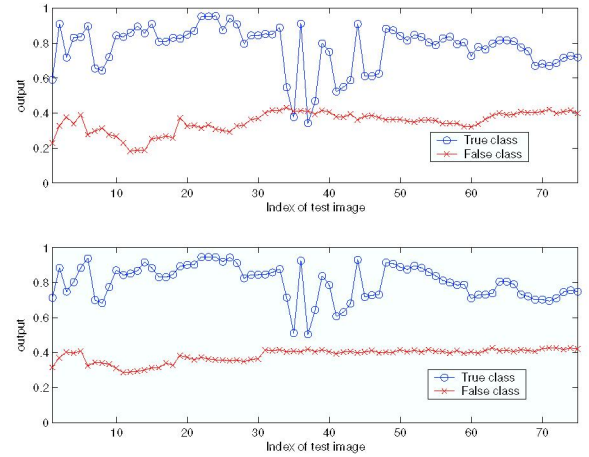


Fig. 2. The averaged test output peak values (100 Monte-Carlo simulations with  $N=5$ ), (Top): Conventional MACE, (Bottom): Correntropy MACE.

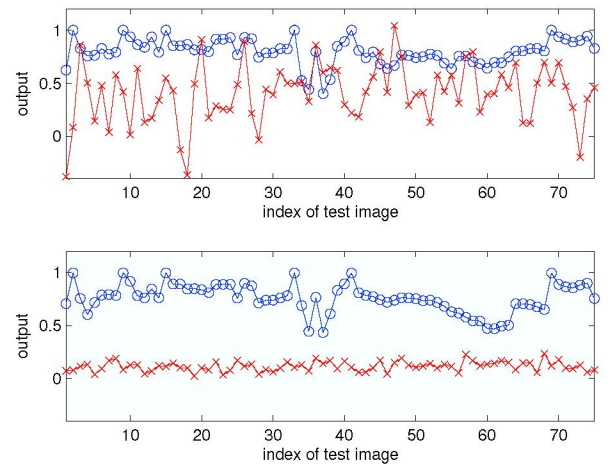


Fig. 3. The test output peak values with additive Gaussian noise ( $N=5$ ), (Top): Conventional MACE, circle-true class with SNR=10dB, cross-false class with SNR=2dB, (Bottom): Correntropy MACE, circle-true class with SNR=10dB, cross-false class with SNR=2dB.



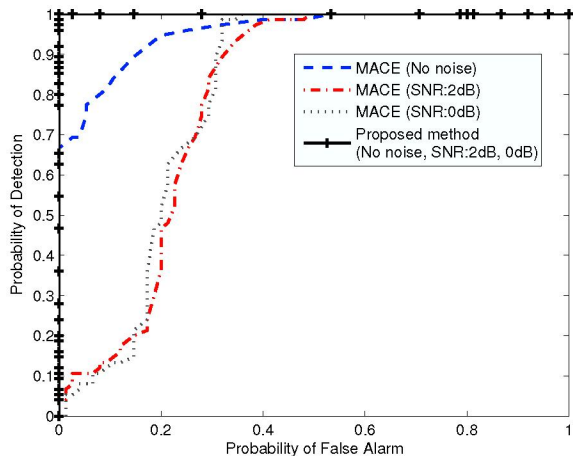


Fig. 4. The comparison of ROC curves with different SNRs.

(true class) and otherwise it should be close to zero. Fig. 2 (Top) shows that the correlation output peak values of the conventional MACE in false classes is close to zero and it means that the MACE has a good rejecting ability of false class. However, some outputs in the test image set, even in the true class, are not recognized as the true class. Fig. 2 (Bottom) shows the output values of the proposed correntropy MACE and we can see that the generalization and rejecting performance are improved. As a result, the two images can be recognized well even with a small number of training images. One of problems of the conventional MACE is that the performance can be easily degraded by additive noise in the test image since the MACE does not have any special mechanism to consider input noise. Therefore, it has a poor rejecting ability for a false class image when noise is added into a false class. Fig. 3 (Top) shows the noise effect on the conventional MACE. When the class images are seriously distorted by additive Gaussian noise (SNR = 2dB), the correlation output peaks of some test images from false class become great than that of the true class, hence wrong recognition happens. The results in Fig. 3 (Bottom) are obtained by the proposed method. The correntropy MACE shows a much better performance especially for rejecting even in a very low SNR environment. Fig. 4 shows the comparison of ROC curves with different SNRs. In the conventional MACE, we can see that the false alarm rate is increased as additive noise power is increased. However, in the proposed method, the probability of detection with zero false alarm rate is 1. The correntropy MACE shows much better recognition performance than the conventional MACE.

One of advantage of the proposed method is that it is more robust than the conventional MACE. That is, the variation of the test output peak value due to a different training set is smaller than that of the MACE. Fig. 5 shows standard deviations of 100 Monte-Carlo outputs per test input when the test

Table 1. Comparison of standard deviations of all the Monte-Carlo simulation outputs (100×75 outputs)

|                      | True<br>(No noise) | False<br>(No noise) | True<br>(SNR:0dB) | False<br>(SNR:0dB) |
|----------------------|--------------------|---------------------|-------------------|--------------------|
| MACE                 | 0.0498             | 0.0086              | 0.0527            | 0.0245             |
| Correntropy-<br>MACE | 0.0488             | 0.0051              | 0.0485            | 0.0038             |

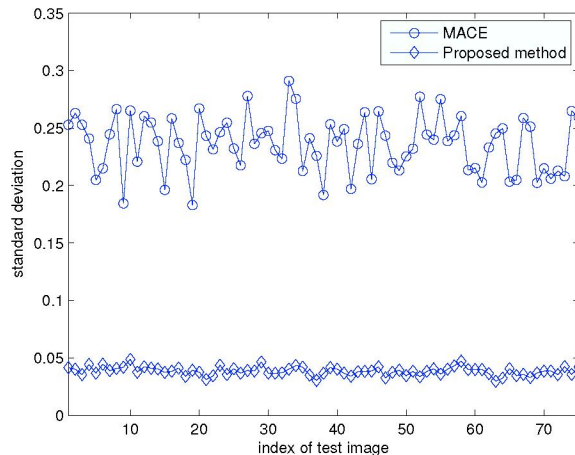


Fig. 5. The comparison of standard deviation of 100 Monte-Carlo simulation outputs of each noisy false class test images.

input are noisy false class images. Table 1 shows the comparison of the standard deviation of 750 outputs (100 Monte-Carlo outputs for 75 inputs) for each class. From the table 1, we can see that the variations of the correntropy MACE outputs due to different training set is much less than those of the conventional MACE and it tells us that our proposed nonlinear version of the MACE outperforms the conventional MACE and achieves a robust performance for distortion-tolerant pattern recognition.

## 5. CONCLUSIONS

In this paper, we have proposed and evaluated a correntropy based nonlinear MACE filter for object recognition. We presented experimental results for face recognition. Using recently introduced correntropy idea, the nonlinear version of the MACE can be implemented in a higher dimensional feature space and this correntropy MACE overcomes the main shortcomings of the conventional MACE which is poor generalization, as well as the effect of the input noise. The correntropy MACE also shows good rejecting performance. This is due to the prewhitening effect in feature space. Simulation results show that the detection and recognition performance of the correntropy MACE is better than that of the MACE in particular in a noisy environment, which indicates that the proposed method is robust and exhibits better distortion tolerance.

rance than the MACE. Although we have shown the simulation results for face recognition, the correntropy MACE filter can be used in other object recognition problems such as the synthetic aperture radar (SAR) ATR system.

## 6. REFERENCES

- [1] B.V.K. Vijaya Kumar, "Tutorial survey of composite filter designs for optical correlators," vol. 31, pp. 4773–4801, 1992.
- [2] M. Savvides, B.V.K. Vijaya Kumar, and P.K. Khosla, "Face verification using correlation filters," in *Proc. Third IEEE Automatic Identification Advanced Technologies*, Tarrytown, NY, 2002, pp. 56–61.
- [3] G.L. Turin, "An introduction to matched filters," *IEEE Trans. Information Theory*, vol. 6, pp. 311–329, 1960.
- [4] A. VanderLugt, "Signal detection by complex spatial filtering," *IEEE Trans. Information Theory*, , no. 10, pp. 139–145, 1964.
- [5] C.F. Hester and D. Casasent, "Multivariant technique for multiclass pattern recognition," *Appl. Opt.*, vol. 19, pp. 1758–1761, 1980.
- [6] B.V.K. Vijaya Kumar, "Minimum variance synthetic discriminant functions," *J. Opt. Soc. Am. A*, vol. 3, no. 10, pp. 1579–1584, 1986.
- [7] A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," *Appl. Opt.*, vol. 26, no. 17, pp. 3633–3640, 1987.
- [8] Ph. Refregier and J. Figue, "Optimal trade-off filter for pattern recognition and their comparison with weiner approach," *Opt. Computer Process.*, vol. 1, pp. 3–10, 1991.
- [9] J.W. Fisher and J.C. Principe, "Recent advances to nonlinear mace filters," *Optical Engineering*, vol. 36, no. 10, pp. 2697–2709, Oct. 1998.
- [10] B. Scholkopf and A. J. Smola, *Learning with Kernels*, The MIT Press, 2002.
- [11] H. Kwon and N. M. Nasrabadi, "Kernel matched signal detectors for hyperspectral target detection," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 2005, vol. 4, pp. 665–668.
- [12] K.H. Jeong, P.P. Pokharel, J. Xu, S. Han, and J.C. Principe, "Kernel synthetic discriminant function for object recognition," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, France, May 2006, vol. 5, pp. 765–768.
- [13] I. Santamaría, P.P. Pokharel, and J.C. Principe, "Generalized correlation function: Definition, properties and application to blind equalization," *IEEE Trans. Signal Processing*, vol. 54, no. 6, pp. 2187–2197, June 2006.
- [14] P.P. Pokharel, R. Agrawal, and J.C. Principe, "Correntropy based matched filtering," in *Proc. IEEE Int. Workshop on Machine Learning for signal Processing (ML-SP)*, Sept. 2005, pp. 148–155.
- [15] P.P. Pokharel, J.W. Xu, D. Erdogmus, and J.C. Principe, "A closed form solution for a nonlinear wiener filter," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, France, May 2006, vol. 3, pp. 720–723.
- [16] "<http://www.amp.ece.cmu.edu> : Advanced multimedia processing lab at electrical and computer eng., cmu," .