Kernel Correlation Filter Based Redundant Class-Dependence Feature Analysis (KCFA) on FRGC2.0 Data

Chunyan Xie, Marios Savvides, and B.V.K. VijayaKumar

Department of Electrical and Computer Engineering,
Carnegie Mellon University, Pittsburgh, PA 15213
{chunyanx, kumar}@ece.cmu.edu, msavvid@ri.cmu.edu

Abstract. In this paper we propose a nonlinear correlation filter using the kernel trick, which can be used for redundant class-dependence feature analysis (CFA) to perform robust face recognition. This approach is evaluated using the Face Recognition Grand Challenge (FRGC) data set. The FRGC contains a large corpus of data and a set of challenging problems. The dataset is divided into training and validation partitions, with the standard still-image training partition consisting of 12,800 images, and the validation partition consisting of 16,028 controlled still images, 8,014 uncontrolled stills, and 4,007 3D scans. We have tested the proposed linear correlation filter and nonlinear correlation filter based CFA method on this FRGC2.0 data. The results show that the CFA method outperforms the baseline algorithm and the newly proposed kernel-based non-linear correlation filters perform even better than linear CFA filters.

1 Introduction

Human face recognition is currently a very active research area [1, 2] with focus on ways to perform robust biometric identification. However, face recognition is a challenging task because of the variability of the appearance of face images even for the same subject as it changes due to expression, occlusion, illumination, pose, aging etc. The Face Recognition Grand Challenge (FRGC) [3] has been organized to facilitate the advancement of face recognition processing across the broad range of topics including pattern recognition algorithm design, sensor design, and in general for advancing the field of face recognition.

In this paper, we focus on the face recognition algorithms based on 2D still images. Many algorithms [4-7] have been developed for face recognition from 2D still images. Among the different approaches, spatial frequency domain methods [8-9] have been shown to exhibit better tolerance to noise and illumination variations than many space domain methods. In this paper, we extend the linear correlation filter to the nonlinear correlation filters using kernel methods. The linear and nonlinear correlation filters are tested on the FRGC2.0 data using the redundant class-dependence feature analysis (CFA) approach. In the CFA method, we train a filter bank of correlation filters based on the data from the generic training set, where we have multiple genuine images for each class. The trained filter bank is then used in

W. Zhao, S. Gong, and X. Tang (Eds.): AMFG 2005, LNCS 3723, pp. 32 – 43, 2005.

validation experiments to extract the discriminant class-dependence features for recognition. The nearest neighbor rule is applied to these features to measure the similarity between target and query images. The algorithm also offers the benefit of computationally efficient training, as when the database size increases there is no need for re-training when a new entry is added to the database.

Kernel tricks have been used with support vector machine (SVM) [10], principal component analysis (PCA) [11], linear discriminant analysis (LDA) [12], kernel spectral matched filter [13] and many other approaches to generate nonlinear classifiers. Motivated by these approaches, we propose in this paper a nonlinear extension of the Equal Correlation Peak Synthetic Discriminant Function (ECP-SDF)[15] filter and the Optimal Trade-off correlation Filter (OTF)[18], obtaining nonlinear correlation filter classifiers for face recognition application. The experimental results show that these nonlinear correlation filters outperform the linear correlation filters in the CFA approach on FRGC2.0 data.

The paper is organized as follows. Section 2 introduces the redundant class-dependence feature analysis method and kernel based nonlinear correlation filters. Section 3 introduces the FRGC2.0 data and Experiments. In Section 4, we show numerical results of the CFA method on the FRGC2.0 data and we discuss the results and outline the future work in Section 5.

2 Redundant Class-Dependence Feature Analysis

Most approaches to face recognition are in the image domain whereas we believe that there are more advantages to work directly in the spatial frequency domain. By going to the spatial frequency domain, image information gets distributed across frequencies providing tolerance to reasonable deviations and also providing graceful degradation against distortions to images (e.g., occlusions) in the spatial domain. Correlation filter technology is a basic tool for frequency domain image processing. In correlation filter methods, normal variations in authentic training images can be accommodated by designing a frequency-domain array (called a correlation filter) that captures the consistent part of training images while de-emphasizing the inconsistent parts (or frequencies). Object recognition is performed by cross-correlating an input image with a designed correlation filter using fast Fourier transforms (FFTs). The advantage of using advanced correlation filter designs is that they offer closed form solutions which are computationally attractive [14].

2.1 Matched Filter (MF) and ECP-SDF Filter

Matched Filters (MFs) [14] are simple correlation filters, which are optimal in the sense that they provide the maximum output signal-to-noise ratio (SNR). However, MFs lose their optimality rapidly when the test image differs from the reference image due to natural variability such as expressions, lighting, pose, etc. For N training images, we need N MFs, one for each training image. The Synthetic Discriminant Function (SDF) approach [15] was proposed to create a composite image that is a linear combination of multiple reference images and the weights for linear combination are selected so that the cross-correlation output at the origin is same for all images belonging to one class. The basic SDF is known as the *equal correlation peak (ECP) SDF* [15]. The objective is to design a composite image **h** such that it generates the same value at the origin of the correlation plane for all training images from the same class. This origin value (loosely referred to as the correlation peak) $c(0,0)$, is the inner product of the training image and the filter to be determined, i.e.,

$$c(0,0) = \mathbf{h}^{+} \cdot \mathbf{x}_i = \mathbf{x}_i^{+} \cdot \mathbf{h} \tag{1}$$

where $\mathbf{x}_i$ denotes the i-th training image and **h** denotes the filter. In most cases, we let $c(0,0)$ be 1 for training images of true class (i.e., authentics) and 0 for the training images of false class (i.e., impostors, assuming that impostor images are available for training). For N training images, we can rewrite (1) as

$$\mathbf{X}^{+}\mathbf{h} = \mathbf{c}^{*} \tag{2}$$

The ECP SDF assumes that the composite image **h** is a linear combination of the training images and it can be solved as in [14]

$$\mathbf{h} = \mathbf{X}(\mathbf{X}^{+}\mathbf{X})^{-1}\mathbf{c}^{*} \tag{3}$$

The ECP SDF filter, however does not incorporate any tolerance to input noise. Also because it is designed solely on the basis of constraints on correlation values at the origin, correlation values elsewhere may be larger and thus the correlation peak may not be the controlled value. If the test input is not centered, then we cannot use it because the correlation output peak is not necessarily the controlled value corresponding to the center of the target. More SDF filters have been developed to address these problems.

2.2 Optimal Tradeoff Filter

Different choices of energy minimization metrics of correlation output lead to correlation filters that address different problems. The *minimum variance synthetic discriminant function* (MVSDF) [16] filter minimizes the correlation output noise energy represented in matrix format as $\mathbf{h}^{+}\mathbf{C}\mathbf{h}$**;** where **C** is a diagonal matrix whose diagonal elements $\mathbf{C}(k,k)$ represent the noise power spectral density at frequency k. The *minimum average correlation energy* (MACE) [17] filter minimizes the average correlation output energy $\mathbf{h}^{+}\mathbf{D}\mathbf{h}$ where **D** is the average of $\mathbf{D}_i$, the power spectrum of the i-th image, which is also a diagonal matrix whose elements $\mathbf{D}_i(k,k)$ contain the power spectra of the i-th training image at frequency k. We note that the MACE filter emphasizes high spatial frequencies in order to produce sharp correlation peaks whereas the MVSDF filter typically suppresses high frequencies in order to achieve noise tolerance. Although both attributes are desired, the corresponding energy metrics cannot be minimized simultaneously. The *optimal tradeoff filter* (OTF) [18] is designed to balance these two criteria by minimizing a weighted metric $\mathbf{h}^{+}\mathbf{T}\mathbf{h}$ where $\mathbf{T}=\alpha\mathbf{D}+\beta\mathbf{C}$ and $0 \le \alpha, \beta \le 1$. The OTF is obtained as shown in (4) below:

$$\mathbf{h}_{OTF} = \mathbf{T}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{T}^{-1}\mathbf{X})^{-1}\mathbf{c}^{*} \tag{4}$$

where $\mathbf{X} = [\ \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\ ]$ is a $d \times N$ matrix, and each $\mathbf{x}_i$ is d dimensional vector constructed by lexicographically reordering the 2-D Fourier transform of the *i*-th training image.

2.3 Redundant Class-Dependence Feature Analysis

When the correlation filters are used for verification, the commonly used method is to correlate the test image with the filter which is designed based on one or more training images, compute PSR value, and to compare it with a preset threshold to decide if the image is authentic or imposter, as shown in Fig. 1.

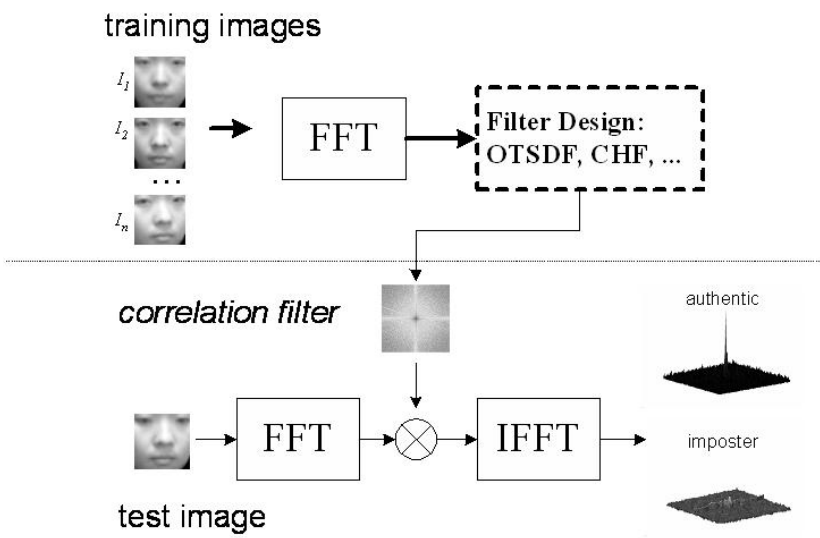


Fig. 1. Commonly used method for still-to-still face verification using correlation filter

There are some problems with this method when applied to the FRGC 2.0 experiments. First, it is not efficient. FRGC2.0 experiment #1 requires that we generate a 16,028x16,028 similarity matrix. For this, we need to design 16,028 correlation filters and compute 16,028x16,028 correlations. It can take a significant amount of time (up to a month using high-power dual processor machines) just to run the whole experiment once. Second, the performance of the filter may not be very good because there is only one genuine image available for training each filter. Third, the generic training set available with FRGC dataset is not being used by the traditional correlation filter synthesis method.

To address these problems, we propose a novel redundant *class-dependence feature analysis* (CFA) method [19] for face recognition using correlation filters. In this method, we train a correlation filter for each subject from the generic training set, and get a bank of subject-dependence correlation filters. All of these filters are used for feature extraction, as shown in Fig. 2. A test image evaluated on all of these filters

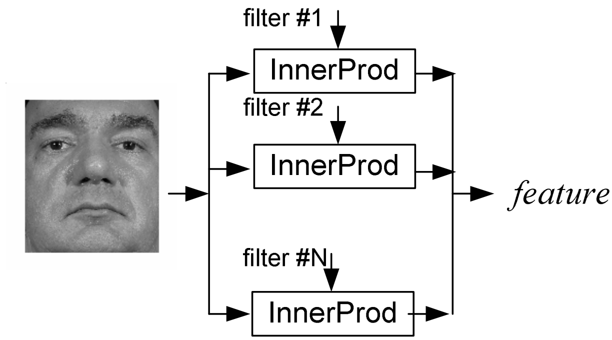


Fig. 2. The concept of feature extraction based on correlation filters

generates a feature vector that is used to represent the test image. Because all of training and test images are centered during the pre-processing stage, we assume that the peak at the correlation output plane is also centered. To make it computationally efficient, we only compute the center value of the correlation output by calculating the inner product of the test image and each synthesized filter. Each component in the feature vector represents the similarity between the test image and a certain subject class. Because all synthetic filters are not orthonormal to each other, the coefficients in the feature vector contain redundant information, so we call this method redundant *class-dependence feature analysis* (CFA).

During the final test matching phase, the nearest neighbor rule is applied to decide the class label for the test image, i.e.,

$$\theta(\mathbf{y}) = \arg_i \min_{i,j} \left(measure\left(\mathbf{t} - \mathbf{r}_{ij} \right) \right) \tag{5}$$

where $\mathbf{r}_{ij}$ represents feature vector corresponding the j-th training sample of the i-th class and $\mathbf{t}$ is the feature vector corresponding to the test input $\mathbf{y}$. There are four commonly used similarity measures: the L_1 norm, the Euclidean distance, the Mahalanobis distance and the cosine function. The cosine distance (6) in our experiments has been shown to be the best performance for this method.

$$S_{\cos}(\mathbf{r}, \mathbf{t}) = \frac{-(\mathbf{r} \cdot \mathbf{t})}{\|\mathbf{r}\| \|\mathbf{t}\|} \tag{6}$$

There are several attributes of the CFA method worth noting. First, when new classes are added in the generic training set, previously trained correlation filters do not require re-training, we just need to add a new filter which is easy to compute due to the nature of the closed form solution of the OTF. Second, since the filter bank is class-dependence, we expect to observe better performance when the generic training set and the validation sets have more overlapped classes. Finally, under the CFA framework, the class-dependence features can also be extracted by some classifiers other than correlation filters, e.g., support vector machines that can be trained for each individual class.

2.4 Kernel Methods of Correlation Filters

Polynomial correlation filter (PCF) [19] had been developed to generate nonlinear correlation filter classifiers. In PCFs some point nonlinearity transforms (e.g. x^2, x^3, etc.) are applied to each pixel of the image and the filters are developed based on the transformed images to optimize a performance criterion of interest. It is shown [19] that the polynomial correlation filter outperforms the linear correlation filter.

In this paper, we introduce a new method to extend the linear ECP-SDF and OTF correlation filters to non-linear correlation filters using kernel methods. As discussed in Sec. 2.3, in face recognition application, we usually assume that the images are centered and geometrically normalized. In that case, we focus on the inner product of the filter and the tested image. For ECP-SDF filter **h** and a test image **y**, we can get

$$c(0,0) = \mathbf{y}^{+}\mathbf{h} = \mathbf{y}\mathbf{X}(\mathbf{X}^{+}\mathbf{X})^{-1}\mathbf{c}^{*} \tag{7}$$

The only way in which the data appears in the correlation framework is in the form of inner products $\mathbf{x}_i \cdot \mathbf{x}_j$. Suppose we map the data to some other feature space by a non-linear mapping Φ, then the correlation peak value of the ECP-SDF filter becomes

$$c(0,0) = \Phi(\mathbf{y}) \cdot \Phi(\mathbf{X})\left(\Phi(\mathbf{X}) \cdot \Phi(\mathbf{X})\right)^{\mathbf{-1}} \mathbf{c}^{*} \tag{8}$$

The training and test algorithms would depend on the functions of the form $\Phi(\mathbf{y}) \cdot \Phi(\mathbf{X})$. Now if we have a kernel function below

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \tag{9}$$

We would only need to use $K(\mathbf{x}_i, \mathbf{x}_j)$ to compute the correlation peak value

$$c(0,0) = K(\mathbf{y}, \mathbf{X})\left(K(\mathbf{X}, \mathbf{X})\right)^{\mathbf{-1}} \mathbf{c}^{*} \tag{10}$$

and we would never need to explicitly know what the Φ mapping is, which saves a lot of computations. This allows us to achieve a nonlinear correlation filter classification boundary in the original image space. Mercer's condition [10] tells us whether or not a prospective kernel is actually an inner product in some space. We use this condition to modify any kernel variations to ensure that this is satisfied.

Next we introduce the method of extend the OTF filter to its nonlinear version. For an OTF filter and a test image **y**, we can get

$$c(0,0) = \mathbf{y}^{+}\mathbf{h} = \mathbf{y}\mathbf{T}^{\mathbf{-1}}\mathbf{X}(\mathbf{X}^{+}\mathbf{T}^{\mathbf{-1}}\mathbf{X})^{\mathbf{-1}}\mathbf{c}^{*} \tag{11}$$

Note that the difference between (11) and (7) is the diagonal matrix **T**, where $\mathbf{T}=\alpha\mathbf{D}+\beta\mathbf{C}$, a linear combination of the input noise power spectral density **C** and average power spectral of the training images. Since **T** is a diagonal and positive definite matrix, it is easy to decompose $\mathbf{T}^{-1} = \mathbf{T}^{-\frac{1}{2}}\mathbf{T}^{-\frac{1}{2}}$, then we can rewrite the correlation peak as

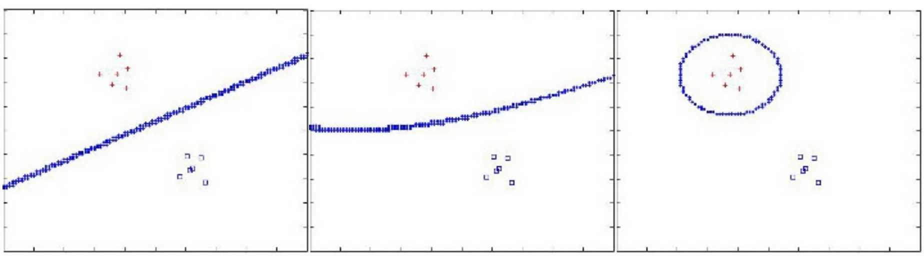
$$c(0,0) = \mathbf{y}\mathbf{T}^{-\frac{1}{2}}\mathbf{T}^{-\frac{1}{2}}\mathbf{X}(\mathbf{X}^{+}\mathbf{T}^{-\frac{1}{2}}\mathbf{T}^{-\frac{1}{2}}\mathbf{X})^{-1}\mathbf{c}^{*} = (\mathbf{T}^{-\frac{1}{2}}\mathbf{y})(\mathbf{T}^{-\frac{1}{2}}\mathbf{X})\left(\left(\mathbf{T}^{-\frac{1}{2}}\mathbf{X}\right)^{+}\left(\mathbf{T}^{-\frac{1}{2}}\mathbf{X}\right)\right)^{-1}\mathbf{c}^{*} \quad (12)$$

We can treat $\mathbf{T}^{-\frac{1}{2}}$ as a pre-processing filter and apply it to every training and test image, so we get

$$c(0,0) = (\mathbf{T}^{-\frac{1}{2}}\mathbf{y})(\mathbf{T}^{-\frac{1}{2}}\mathbf{X})\left(\left(\mathbf{T}^{-\frac{1}{2}}\mathbf{X}\right)^{+}\left(\mathbf{T}^{-\frac{1}{2}}\mathbf{X}\right)\right)^{-1}\mathbf{c}^{*} = \mathbf{y}'\mathbf{X}'\left(\mathbf{X}'^{+}\mathbf{X}'\right)^{-1}\mathbf{c}^{*} \quad (13)$$

which is in the same form as in (7) and we can apply the kernel trick as well to obtain the kernel based nonlinear OTF classifier. In Fig. 3 we show an original image and a pre-filtered image by **T** with α=0.1 and β =1.

Fig. 3. Illustration of image preprocessing: (a) original image and (b) pre-filtered image



(a) (b) (c)

Fig. 4. The illustration of the decision boundary for (a) linear correlation filter, (b) polynomial correlation filter and (c) Gaussian RBF classifier

Some useful kernels include the polynomial kernel in (14) which results in a classifier that is a polynomial of degree *p* in the data and the Gaussian radial basis function (RBF) kernel in (15) that gives a Gaussian RBF classifier. In Fig. 4 we use a toy example to show the linear correlation filter and polynomial correlation filter and Gaussian RBF classification boundary.

$$K(\mathbf{x},\mathbf{y}) = (\mathbf{x}\cdot\mathbf{y}+1)^p \quad (14)$$

$$K(\mathbf{x},\mathbf{y}) = e^{-\|\mathbf{x}-\mathbf{y}\|^2/2\sigma^2} \quad (15)$$

3 FRGC 2.0 Data and Experiments

For the second phase of the face recognition grand challenge FRGC [3] data 2.0 was again collected at the University of Notre Dame. The generic training set contains 222 subjects and consists of 12,776 still images captured in 100 subject sessions from academic year 2002-2003, and the validation set contains 466 subjects and consists of 16,028 controlled still images and 8,014 uncontrolled still images captured in 4007 subject sessions from academic year 2003-2004. The example images from the controlled still image set are shown in Fig. 5. The FRGC experiment #1 is defined to generate a similarity matrix of 16,028x16,028 similarity scores of controlled indoor still images vs. indoor still images and the FRGC experiment #4 is defined to generate a 16,028x8,014 similarity matrix of controlled still images vs. uncontrolled still images. Experiment #1 data set only exhibits some facial expression, minimal illumination variations and minimal pose variations and Experiment #4 data contains severe illumination variations and blurring. Experiment #4 is much harder because the

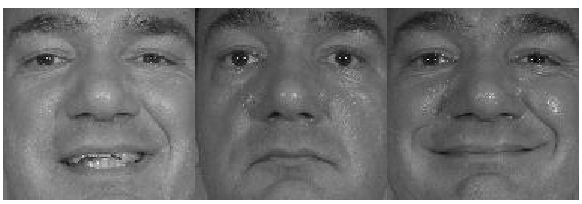


Fig. 5. Example images from the controlled still set

Fig. 6. Example images with illumination variations and blurring from the uncontrolled still set

query images are of poorer quality. More details of all experiments defined in FRGC project can be found in [3].

Fig. 6 shows example images from uncontrolled still set that contains images under severe illumination conditions and images out of focus. We can see that the combination of the illumination, expression, pose variations and the blurring effect makes the recognition task even harder.

4 FRGC Numerical Results

In this paper, we focus on the FRGC2.0 experiment #1 for which we test our proposed algorithms on a controlled 2D still image set of 16,028 images and generate a similarity matrix of size 16,028x16,028. In this paper we show the experimental results of five different algorithms using our proposed CFA approach. The first one is CFA based on the OTF correlation filter with parameters $\alpha=0.001$ and $\beta=1$. The second one is the nonlinear CFA based on the polynomial kernel with polynomial degree parameter $p=2.0$. The third one is the nonlinear CFA based on the fractional power polynomial model [20] with polynomial degree parameter $p=0.9$. Note that when $p=0.9$, a fractional power polynomial does not necessarily define a kernel function, as it might not define a positive semi-definite Gram matrix [20] but the fractional power polynomial models with $0<p<1$ shows better face recognition performance than polynomial kernel with $p>1$ [20] using the same fractional power trick. It is also observed in our experiments. The forth one is the nonlinear CFA based on Gaussian RBF classifier with the variance parameter $\sigma^2 = 3.0$. The last one is also the nonlinear CFA based on Gaussian RBF classifier ($\sigma^2 = 3.0$) plus a preprocessing filter **T** ($\alpha=0.1$ and $\beta=1$).

In Fig. 7, we show three OTF correlation filter bases trained on the training data for the linear CFA approach. We generate the CFA feature vectors for all target images and query images by inner product and apply the nearest neighbor rule based on cosine distance to get the similarity matrix.

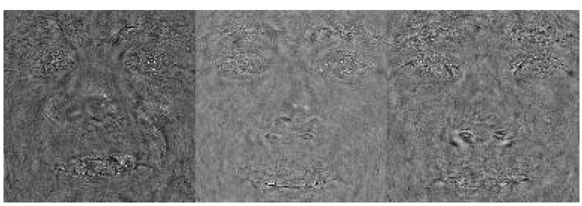


Fig. 7. Three correlation filter-based CFA basis

For the nonlinear CFA, we use the corresponding kernel functions or models with above specified parameters. We show the verification performance of 5 different algorithms on the FRGC2.0 experiment #1 in table 1 and Fig. 8. In table 1, we reported the face verification rate when the false acceptance rate (FAR) equal to 0.1% which are the result specifications according to FRGC. We compare results of the five algorithms described above to the FRGC baseline result and the traditional correlation filter results.

Table 1. The verification rate at 0.1% FAR of the different methods

Baseline PCA	OTF	CFA Linear Filter	CFA Poly (0.9)	CFA Poly (2.0)	CFA RBF	CFA RBF+OTF
66% [3]	77%	89.9%	91.4%	90.2%	92.5%	93.5%

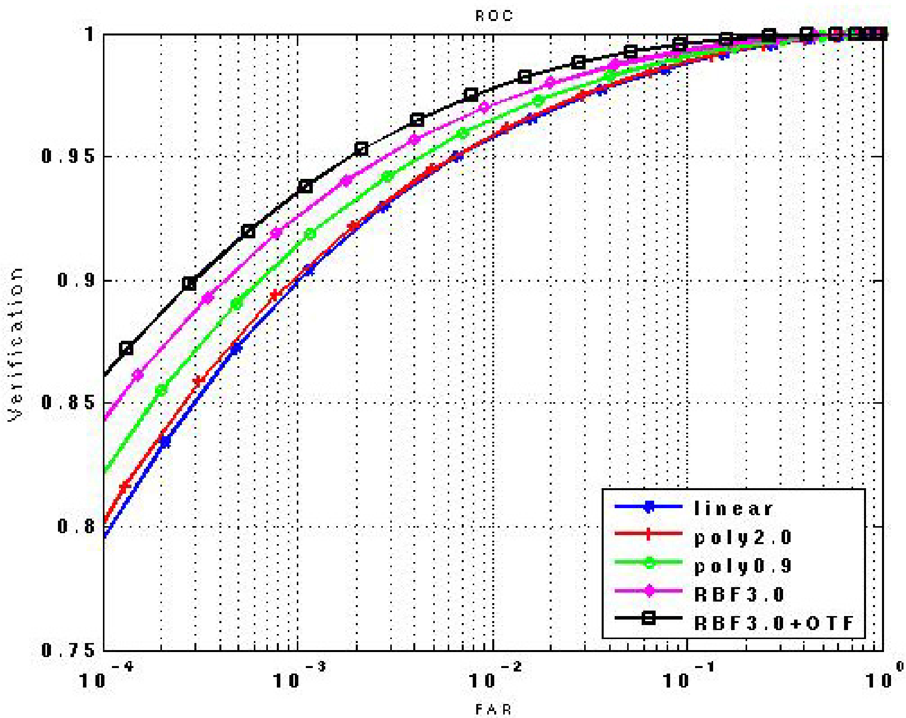


Fig. 8. The ROC curves of the verification performance of 5 different algorithms on FRGC2.0 experiment #1

From table 1 we can see that the CFA approach perform much better than the baseline PCA algorithm and it also performs better than the traditional method of face verification using correlation filter. The nonlinear CFA method generally outperforms the linear CFA method, and the kernel method plus the optimal tradeoff filter perform the best. The ROC curves of these experimental results are shown in Fig. 8.

From Fig. 8 we can see that the nonlinear correlation filters based CFA approach generally outperforms linear correlation filters. The Gaussian RBF classifier based CFA method significantly improves the verification performance over other methods. Moreover, the Gaussian RBF plus the OTF based pre-processing filter further improves the performance; clearly showing the advantages of frequency domain approaches. From our experiments we can also see that there are some parameters to be decided when apply the nonlinear kernel method and/or the linear correlation filters. In this paper, we show the best results of testing on different parameters in our experiments. In our future work, we will investigate the methods on how to theoretically and experimentally select the nonlinear parameter and the correlation filter parameters. We will also investigate other possible kernel functions for nonlinear approaches.

5 Conclusions

In this paper we introduce a novel kernel correlation filter method applied in the redundant class-dependence feature analysis (CFA) approach to perform robust face recognition in the Face Recognition Grand Challenge (FRGC) data set. Under the CFA framework, the class-dependence features can be extracted for each target image and the query image and the nearest neighbor rule is applied for classification. The linear correlation filter based CFA approach is extended to the nonlinear correlation filter based CFA approach with the kernel methods. The verification performance of the linear filter and nonlinear filter based CFA approaches have been tested on the FRGC2.0 experiment#1, and the verification rates (93.5%) are much better than that reported of the baseline algorithm (66%). In the future, we will theoretically and experimentally investigate the method of how to select the kernel function parameters and correlation filter parameters and aim to further improve the face verification performance on the FRGC experiments. We will also investigate the CFA approaches based on other linear and non-linear classifiers (e.g. support vector machine) and extend other different types of advanced correlation filters to kernel based nonlinear filters.

Acknowledgements

This research is sponsored by the Technical Support Working Group (TSWG) and also in part by Carnegie Mellon University's CyLab.

References

1. R. Chellappa, C.L. Wilson, S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE,* Vol.83 (5), pp705-741 (1995).
2. R. Gross, J. Shi, and J. Cohn, "Quo Vadis Face Recognition?" *Third Workshop on Empirical Evaluation Methods in Computer Vision*, December 2001.
3. P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the Face Recognition Grand Challenge," *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
4. M. Turk, A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, 3(1), pp71-86, 1991.
5. A. Pentland, B. Moghaddam, T. Starner, "View-Based and Modular Eigenspaces for Face Recognition," *IEEE CVPR,* 1994.
6. P.Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. PAMI*-19(7),(1997).
7. W. Zheng, C Zou, L. Zhao, "Real-time face recognition using Gram-Schmidt orthogonalization for LDA," *ICPR'04*, August, 2004, Cambridge UK, pp. 403-406

8. B.V.K. Vijaya Kumar, M. Savvides, K. Venkataramani, and C. Xie, "Spatial Frequency Domain Image Processing For Biometric Recognition," P*roceedings IEEE International Conference on Image Processing*, pp. 53-56, 2002
9. M. Savvides, B.V.K. Vijaya Kumar, and P.K. Khosla, "'Corefaces'- Robust Shift Invariant PCA based correlation filter for illumination tolerant face recognition," *Proceedings IEEE Computer Vision and Pattern Recognition (CVPR)*, pp. 834-841, June 2004.
10. V. Vapnik, *The nature of statistical learning theory*, Springer-Verlag, New York, 1995.
11. Schölkopf, A. Smola, and K.-R. Müller. "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, 10:1299-1319, 1998
12. G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach,",*Neural Computation*, vol. 12, pp. 2385–2404, 2000.
13. N. M. Nasrabadi and H. Kwon, "Kernel spectral matched filter for hyperspectral target detection," *Proceedings ICASSP*, 665-668., 2005
14. B.V.K. Vijaya Kumar, "Tutorial survey of composite filter designs for optical correlators," *Appl. Opt*., 31, 4773-4801, (1992)
15. C.F. Hester and D. Casasant, "Multivariant technique for multiclass pattern recognition," *Appl. Opt.* 21, 4016-4019 (1982)
16. B.V.K. Vijaya Kumar, "Minimum Variance synthetic discriminant functions," *J. Opt. Soc. Am A*, Vol 3, pp1579-1584.
17. A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," *Appl. Opt*. 26, pp. 3633-3630 (1987)
18. P. Refregier, "Filter Design for optical pattern recognition: Multi-criteria optimization approach," *Opt. Lett.*, V.15, 854-856, 1990.
19. Chunyan Xie, Marios Savvides and B.V.K. Vijaya Kumar, "Redundant Class-Dependence Feature Analysis Based on Correlation Filters Using FRGC2.0 Data," *IEEE Workshop on Face Recognition Grand Challenge Experiments in conjunction with CVPR 2005*, San Diego, CA, 21 June 2005
20. B. V. K. Vijaya Kumar and A. Mahalanobis, "Recent advances in composite correlation filter designs," *Asian Joural of Physics*, Vol. 8, No. 4, pp. 407-420, 1999
21. C. Liu, "Gabor-based Kernel PCA with Fractional Power Polynomial Models for Face Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 572-581, 2004