Generalized Low Dimensional Feature Subspace for Robust Face Recognition on Unseen datasets using Kernel Correlation Feature Analysis

Ramzi Abiantun, Marios Savvides, and B.V.K. Vijayakumar
ECE Department, Carnegie Mellon University, USA
{raa@andrew, msavvid@ri, kumar@ece}.cmu.edu

ABSTRACT

In this paper we analyze and demonstrate the subspace generalization power of the kernel correlation feature analysis (KCFA) method for producing compact low dimensional subspace that has good representation ability to work on unseen, untrained datasets. Examining the portability of an algorithm across different datasets is an important practical aspect of face recognition applications where the technology cannot be dataset-dependant in real-world practical applications. In most face recognition literature, algorithms are demonstrated on datasets by training on some part of the dataset and testing on the remainder. In general, the training and testing data have the same people but different capture sessions so essentially, some of the expected variation and people are modeled in the training set. In this paper we describe how we efficiently build a compact feature space using kernel correlation filter analysis on the generic training set of the FRGC dataset, and test the built subspace on other well-known face datasets. We show that the feature subspace produced by KCFA has good representation and discrimination to unseen datasets and produces good verification and identification rates compared to other subspace methods such as PCA. Its efficiency, lower dimensionality (the KCFA is only a 222 dimensional subspace) and discriminative power make it more practical and powerful than PCA as a powerful lower dimensionality reduction method for modeling faces and facial variations.

Index Terms**—** Reduced Feature Subspace, Kernel Correlation Filters, FRGC, PIE, FERET

1. INTRODUCTION

Face recognition continues to be a very popular and active research area due to the increasing demand for access control and surveillance watch-list applications. This research has been made possible by the numerous datasets of face pictures gathered over the years and made available to the face recognition community [1]. Most of these datasets were designed with a specific purpose in mind, in order to advance and test the robustness of algorithms to specific challenges and variations, such as head pose, expression, lighting conditions, surrounding settings and camera angle. However, the process of developing an algorithm based on a particular dataset makes it inherently dependent on it, no matter how large and diversified the given dataset is. Moreover, overtraining becomes an issue, in a way that an algorithm that performs very well on a given dataset is not guaranteed to perform well on others. In this paper we study how our Kernel Correlation Feature Analysis (KCFA) algorithm performs when applied across different face databases. We trained on the FRGCv2 generic set to extract a feature subspace of 222 dimensions only. When tested on the FRGC Experiment 4 (the hardest experiment), the verification rate of KCFA exceeds 82% at 0.1% FAR, while Principal Component Analysis (PCA) [2], the baseline algorithm in FRGC experiments, yields a verification rate of 12% at 0.1 % FAR. We demonstrate that this performance gain carries over to other unseen datasets, proving that our low dimensional feature subspace we built from FRGC generalizes to other unseen datasets better than traditional subspace modeling such as PCA.

2. KERNEL CORRELATION FEATURE ANALYSIS

The traditional face techniques based on dimensionality reduction such as PCA [2], LDA [3] and variants [4] have been proved to under-perform in face recognition application, due to the highly nonlinear nature of face distortions. Correlation filters, with advantages such as noise tolerance, shift-invariance and closed-form solutions to optimization criteria represent a better technique.

2.1 Correlation Filter Theory

The basic correlation filter is the synthetic discriminate function (SDF) [5], a linear combination of training images that is designed to produce a correlation output with a preset value at the origin. An evolution of the SDF is the minimum average correlation energy (MACE) [6] [10] filter. It is designed to minimize the average correlation plane energy resulting from the training images, while constraining the correlation peak value at the origin to pre-specified values. Correlation outputs from MACE filters typically exhibit sharp peaks, making the peak detection and location relatively easy and robust. The closed form expression for the vectorized MACE filter **h** is

$$\mathbf{h} = \mathbf{D}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{D}^{-1}\mathbf{X})^{-1}\mathbf{u} \tag{1}$$

where **X** is a d^2xN complex valued matrix. Here N is the number of training images and d^2 is the number of pixels in each image. Each column contains the lexicographically re-ordered version of the 2-D Fourier transform of a certain training image**.** **D** is a d^2xd^2 diagonal matrix containing the average power spectrum of the training images along its diagonal and **u** is a column vector containing N pre-specified correlation values at the origin.

In the case where the images have been corrupted by additive zero-mean stationary noise with power spectral density represented by **C** (a diagonal matrix whose diagonal elements represent the noise power spectral density at a given frequency), we can trade off some peak sharpness for some noise tolerance using the optimal trade-off filter (OTF) [7] given by:

$$\mathbf{h} = \mathbf{P}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{P}^{-1}\mathbf{X})^{-1}\mathbf{u} \tag{2}$$

Where $\mathbf{P} = \alpha\mathbf{D} + \sqrt{(1-\alpha^2)}\mathbf{C}$ and $0 \leq \alpha \leq 1$

These correlation filters have been shown to exhibit robustness to illumination variations and other distortions [5] [9]. Note that in the formulation of the correlation filters mentioned above, **X** can include impostor training images and **u** can set their corresponding constraint values to 0. Our class-dependent feature analysis (CFA) makes use of the generic dataset to build a set of filters as we explain next. This way we achieve a novel dimensionality reduction far better than the traditional PCA.

2.2 Class Dependent Feature Analysis (CFA)

Our proposed method uses a set of MACE filters to extract features from the training set. For every subject in the training set, we build a MACE filter to end up with as many different filters as there are individuals in the training set. Each filter takes as input all of the training face images available. As noted earlier, the **u** vector in equations (1) and (2) presets the value of the correlation peak. In particular, for the "authentic" class to whom the MACE filter belongs, the **u** values of all pictures belonging to this authentic class are set to 1. While for all other images belonging to the remaining "impostor" classes, we set the equivalent **u** values to 0. This ensures that the filter finds no correlation between subjects belonging to different classes. Therefore, the proposed CFA method is a function of the number of classes and not the total number of training images, which is beneficial for dimensionality reduction in face recognition applications.

Once the design process is completed the testing feature extraction is the next step where we are given a test image **y**, we represent it by the correlation of that test image with the N MACE filters:

$$\mathbf{c} = \mathbf{H}^{\mathrm{T}}\mathbf{y} = [\mathbf{h}_{\mathbf{mace\text{-}1}}\ \mathbf{h}_{\mathbf{mace\text{-}2}} \ldots\ \mathbf{h}_{\mathbf{mace\text{-}n}}]^{\mathrm{T}}\mathbf{y} \tag{3}$$

$\mathbf{h}_{\mathbf{mace}\text{-}i}$ is a filter that is trained to give a small correlation output (close to 0) for all classes except for class-i as shown in Figure 1. Then each input image **y** is projected onto those basis vectors to yield an N-dimensional correlation feature vector **c**, where N is the number of training subjects.

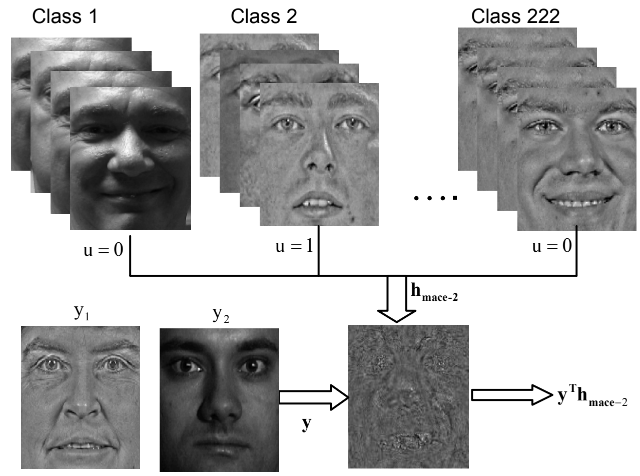

Figure 1: The CFA algorithm; in this figure, we are building a MACE filter for class 2. Note that **u** corresponding to all other classes is set to 0. We are testing pictures $\mathbf{y_1}$ and $\mathbf{y_2}$. The filter response of $\mathbf{y_1}$ and $\mathbf{h}_{\mathbf{mace\text{-}2}}$ can be distinctive to that of $\mathbf{y_2}$ and $\mathbf{h}_{\mathbf{mace\text{-}2}}$

2.3 Kernel Correlation Filters

The linear subspace approach may not perform well, especially if the training data is not well representative, due to the nonlinear distortions in human face appearance variations. To overcome this hurdle, our algorithm is extended to represent nonlinear features efficiently by mapping onto a higher dimensional feature space. This will allow us to exploit higher-order correlations in these kernel spaces. Moreover, to keep the computation tractable even with the increase in dimensionality, we revert to the kernel trick methods for improved efficiency. They enable us to obtain the inner products in the higher-dimensional feature space without actually having to form the higher-dimensional feature mappings. Examples are Kernel Eigenfaces and Kernel Fisherfaces [12]. The mapping function can be denoted as follows.

$$\Phi : R^N \rightarrow F \tag{4}$$

Kernel functions defined by $K(\mathbf{x},\mathbf{y}) = <\Phi(\mathbf{x}), \Phi(\mathbf{y})>$ can be used without having to form the mapping as long as kernels form an inner-product and satisfy Mercer' theorem to ensure that we are still working in a Hilbert inner-product space. In this paper we use a Radial Basis Function kernel whose equation is given by:

$$K(\mathbf{a},\mathbf{b}) = \exp(-(\mathbf{a}-\mathbf{b})^2 / 2\sigma^2) \tag{5}$$

Now we can apply the kernel trick to yield the Kernel Correlation Filter as follows below:

$$\begin{aligned} \Phi(\mathbf{y})\cdot\Phi(\mathbf{h}) &= (\Phi(\mathbf{y})\cdot\Phi(\mathbf{X}'))(\Phi(\mathbf{X}')\cdot\Phi(\mathbf{X}'))^{-1}\mathbf{u} \\ &= K(\mathbf{y},\mathbf{x}'_i)K(\mathbf{x}'_i,\mathbf{x}'_j)^{-1}\mathbf{u} \end{aligned} \tag{6}$$

We can use these Kernel filters in the same CFA framework to extract *N* dimensional kernel feature vectors, which we refer to as the KCFA method.

3. DATASETS

To apply this algorithm in this paper, we chose the FRGC generic set for training because it contains the most face images with the most variations to date. We tested on the biggest other databases available, namely PIE, AR and FERET. Note that we compensate for the illumination variations by preprocessing all of our training and testing pictures [13].

3.1 FRGC

The Face Recognition Grand Challenge (FRGC) database contains over 36000 pictures [14]. In this paper, we use the "generic dataset" which contains 222 people and a total of 12,776 training facial images. This training set features both indoor/controlled and outdoor/uncontrolled pictures with harsh illumination variations (Figure 2). Note that none of the people in the FRGC dataset are present in AR, PIE or FERET.

3.2 Testing Datasets (PIE, AR, FERET)

The CMU Pose, Illumination and Expression (PIE) database [15] contains over 40,000 facial images of 68 different subjects at different poses, with different illumination variations and shadow artifacts. In this paper, we use two different subsets of PIE; as gallery set, we use PIE "light", which features illumination variations while ceiling lights are on, and as probe set, we use PIE "no light", which contains the same subjects with significantly harsher illumination variations due to the fact that ceiling lights were turned off during capture. PIE "light" has 1584 images while PIE "no light" contains 1386 images.

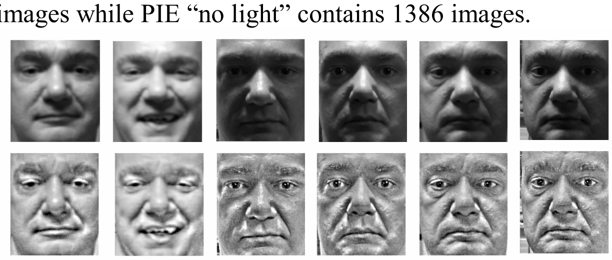

Figure 2: FRGC uncontrolled images before and after preprocessing for illumination variation

The Face Recognition Technology [FERET] dataset [16], one of the most widely used benchmark databases, features facial images in a semi-controlled environment. The data we used for testing contains 599 classes and a total of 1708 images. Similarly, the AR database [17] contains face images of 135 subjects captured over two sessions. For this paper, we're using a subset of 1771 controlled face images that contains no visual obstructions such as sunglasses and scarves.

4. EXPERIMENTAL RESULTS

As detailed above, we built our lower dimensional feature subspace on the FRGC generic training set and project images of other datasets on this reduced dimensional feature space. We compare our KCFA subspace to using actual raw data from the testing dataset images themselves using nearest neighbor distance classifier (with cosine distance measure). We also compare to it to PCA using 1000 and 222 eigenfaces trained on the same FRGC generic set. Normalized cosine distance was the metric used to measure similarity between two images.

Table 1 shows the rank 1 identification rates and table 2 presents the total accept rates taken at 0.1% FAR.

Dataset	Rank 1 Identification Rates			
	KCFA 222 features	Norm correlation	PCA 1000 eigenfaces	PCA 222 eigenfaces
PIE sets	100	100	99.64	99.57
FERET	89.64	33.49	63.93	60.19
ARDB	93.62	78.15	87.80	84.87

Table 2: Verification Rates (VR) at 0.1% FAR

Dataset	Verification Rates @ 0% False Acceptance Rate			
	KCFA 222 features	Norm correlation	PCA 1000 eigenfaces	PCA 222 eigenfaces
PIE sets	56.17	02.30	03.93	04.25
FERET	73.14	30.86	26.91	26.98
ARDB	46.59	20.15	35.08	33.29

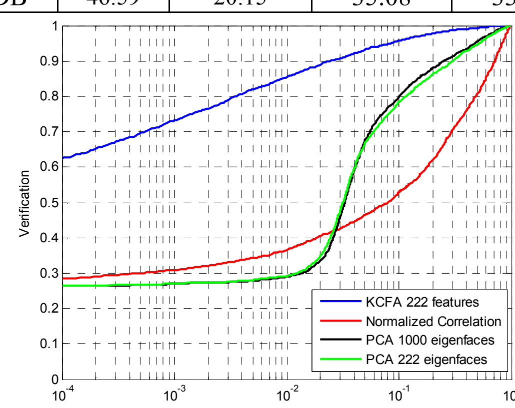

Figure 4: FERET VR vs. FAR using KCFA, Normalized Correlation, PCA 1000 and 222 eigenfaces.

Figures 4 above and 5 below depict the ROC curves for the PIE and FERET databases using the four methods described above.

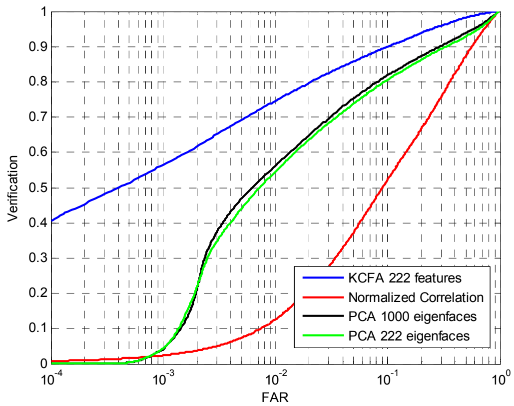

Figure 5: PIE light vs. nolight: VR vs. FAR using KCFA, normalized Correlation, and PCA (1000 and 222 eigenfaces).

Note that the 222 number corresponds to the total number of generic training subjects. By excluding some of these subjects, starting with the ones with the fewest pictures, we can further reduce the size of our feature subspace. Figure 6 below depicts how rank-1 identification rates decay with the smaller number of KCFA features.

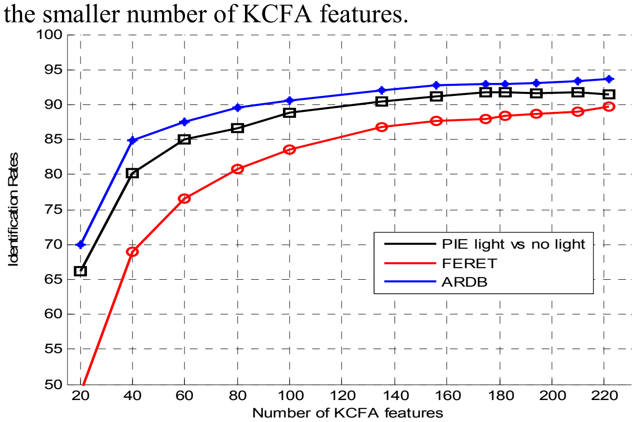

Figure 6: Identification Rates for the different experiments with varying number of KCFA features.

5. CONCLUSIONS

Linear approaches such as PCA, LDA, and CFA may not be suitable to represent or discriminate facial features efficiently. We showed that the low dimensional feature subspace produced by KCFA has good representation and discrimination and generalization to unseen datasets and produces better verification and identification rates on PIE, FERET and AR dataset compared to PCA and also outperforms using raw gallery data The proposed approach is a very efficient dimensionality reduction method reducing the representation of the 12,776 FRGC generic facial images to a KCFA feature subspace of size 222. This allows for efficient transmission on low bandwidth connections such as cell phones and PDA's and permits for very fast matching and low search times in large databases due to the small feature size representation.

6. ACKNOWLEDGEMENTS

This research was sponsored in part by the United States Technical Support Working Group (TSWG) (special thanks to Dr. David Herrington and Dr. Jim Wayman) and in part by Carnegie Mellon CyLab.

REFERENCES

[1] R. Gross, "Face Databases," *Handbook of Face Recognition*, S.Li, A.Jain, ed., Springer, New York, February, 2005.

[2] M. Turk and A. Pentland, "Eigenfaces for Recognition," Journal of *Cognitive Neuroscience*, Vol. 3, pp.72-86, 1991.

[3] P.Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. PAMI*, Vol19 No.7 pp.711-720, 97.

[4] L.F. Chen, H.Y.M. Liao, M.T. Ko, J.C. Lin, and G.J. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Pattern Recognition*, Vol. 33, pp. 1713-1726, 2000

[5] D. Casasent, *Unified Synthetic Discriminant Function Computational Formalution*, Applied Optics, Vol. 23, No. 10, pp. 1620-1627, May 1984.

[6] A. Mahalanobis, B.V.K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," *Appl. Opt*. 26, pp. 3633-3630, 1987.

[7] B.V.K Vijaya Kumar, D.W. Carlson and A. Mahalanobis, *Optimal tradeoff synthetic discriminant function filters for arbitrary devices,* Optics Letters, Vol. 19 pp 1556-1558, 94

[8] C. Xie, M. Savvides, and B.V.K. Vijaya Kumar, "Redundant Class-Dependence Feature Analysis Based on Correlation Filters Using FRGC2.0 Data," *IEEE Conf. Computer Vision and Pattern Recognition(CVPR)* , 2005

[9] M. Savvides, B.V.K. Vijaya Kumar and P.K. Khosla, "Face verification using correlation filters," *Proc. Of Third IEEE Automatic Identification Advanced Technologies*, Tarrytown, NY, pp.56-61, 2002.

[10] M.H Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition using Kernel Methods," *IEEE Conf. Automatic Face and Gesture Recognition*, pp. 215-220, 2002

[11] K. Fukunaga, *Introduction to Statistical Pattern Recognition* (2nd Edition), *New York: Academic Press*, 1990

[12] V. N. Vapnik, *the Nature of Statistical Learning Theory*, New York: Springer-Verlag, 1995.

[13] R Gross and V. Brajovic, "An Image Preprocessing Algorithm for Illumination Invariant Face Recognition", *4th International Conf. on Audio and Video Based Biometric Person Authentication (AVBPA),* 2003

[14] 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," *IEEE Conf. Comp Vision and Pattern Recognition(CVPR)*, 2005

[15] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 12, December, 2003, pp. 1615 - 1618.

[16] J. Phillips, H Moon, S Rizvi, P Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, October 2000.

[17] A.R. Martinez and R. Benavente. The AR face database. Technical Report 24, Computer Vision Center (CVC) Technical Report, Barcelona, 1998