Cafing Shan

Capturing Canonical Correlations for Facial Analysis

Paper No.:

Canonical Correlation Analysis (CCA) is a statistical technique developed by H. Hotelling [1] for measuring linear relationships between two multidimensional variables. It finds base vectors (canonical factors) for two variables such that the correlations between the projections of the variables onto these canonical factors are mutually maximized. The directions of canonical factors capture functional relations of the two variables. Canonical correlations are also known as cosines of principal angles between two linear subspaces [2]. Recently CCA has been applied to computer vision and pattern recognition problems [3, 4, 5, 6, 7, 8].

CCA can be used to establish the relation between two sets of measurements. Borga [3] adopted CCA and phase analysis to find corresponding points in stereo images. Melzer [4] applied CCA to model the relation between an object's poses with raw brightness images for appearance-based 3D pose estimation. Harsoon [5] presented a method using CCA to learn a semantic representation to web images and their associated text.

Like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), CCA also reduces the dimensionality of the original variables, since only a few factor pairs are normally needed to represent the relevant information. However, they serve different purposes: whilst PCA aims to minimize the reconstruction error, and LDA derives a discriminant function that maximizes between-class scatter and minimize within-class scatter, CCA seeks directions for two variables to maximize their correlations, so it is better suited for regression tasks. Compared to other linear regression methods such as Partial Least Squares and Multivariate Linear Regression, CCA has some attractive properties. For example, CCA is invariant to affine transformations of the input variables [7]. Recently Donner [7] presented a fast Active Appearance Model search algorithm based on CCA, which uses reduce-rank regression estimates obtained by CCA, instead of standard linear least-square regression estimates. The CCA based method extracts additional regression-relevant information which may be discarded by the PCA-based model. Reiter [8] proposed to predict 3D depth maps of faces and near-infrared face texture from color face images using CCA.

Many computer vision problems can be cast as learning problems over image sets, for example, object recognition, where a set of images represents the variation in an object's appearance. Recently CCA has attracted increasing attention for image set matching [9, 10, 6], where each set is represented by a linear subspace and the principal angles between two subspaces are exploited as a similarity measure. As a way of comparing sets of images, canonical correlations offer many benefits in accuracy, efficiency, and robustness compared to the classical parametric distribution-based and non-parametric sample-based methods [9]. Recently Kim [6] developed a discriminative learning method over sets for set classification, which maximizes the canonical correlations of within-class sets and minimizes the canonical correlations of between-class sets.

Existing work using CCA on image data is, without exception, based on the vector-space model, that is, the original two-dimensional images are reshaped into one-dimensional vectors, and the collection of image data is modeled as a single data matrix. However, this matrix-to-vector operation leads to two main problems. Firstly, the intrinsic 2D structure of an image matrix is removed, so the spatial information stored therein is discarded. CCA based on this model can not fully capture correlations between 2D image data. Secondly, each image sample is modeled as a high-dimensional vector so that a large number of training samples are needed to yield a reliable and robust estimation of the underlying data distribution. However, in reality, very limited number of data are usually available.

To address these problems, we introduce in this paper a Two-Dimensional Canonical Correlation Analysis (2DCCA) for correlation analysis of 2D image or matrix data. 2DCCA takes a 2D matrix based data representation rather than the 1D vector based representation in classical CCA. So the collection of data is represented as a collection of matrices, instead of a single large matrix. Unlike classical CCA, there is no closed form solution for the optimization problem of 2DCCA. Instead, we propose an iterative algorithm to compute 2-dimensional canonical factors and canonical correlations. We extensively evaluate the proposed 2DCCA in establishing the relation between image sets for regression based facial analysis. Experimental results demonstrate 2DCCA can better capture and measure correlations in 2D images data, providing superior performance in regression and recognition tasks, whilst in the mean time requiring much fewer canonical factors, resulting in lower computational cost. Fig. 1 shows some facial part analysis results. The underlying reason is that 2DCCA is able to preserve and utilize the intrinsic 2D spatial structure in image data.

Recently Two-Dimensional PCA (2DPCA) [11, 12, 13] and Two-Dimensional LDA (2DLDA) [14, 15] have been proposed to improve the traditional 1D vector based PCA and LDA. However, as 2DPCA aims to minimize the reconstruction error and 2DLDA seeks to derive a discriminant function, none of them is best suited for measuring relationships of two matrix/image variables.

References

[1] H. Hotelling. Relations between two sets of variates. *Biometrika*, 8:321–377, 1936.

[2] G. H. Golub and H. Zha. The canonical correlations of matrix pairs and their numerical computation. Technical report, 1992.

[3] M. Borga. *Learning Multidimensional Signal Processing*. PhD thesis, Linkoping University, SE-581 83 Linkoping, Sweden, 1998. Dissertation No 531.

[4] T. Melzer, M. Reiter, and H. Bischof. Appearance models based on kernel canonical correlation analysis. *Pattern Recognition*, 39(9):1961–1973, 2003.

[5] D. Hardoon, S. Szedmak, and J. Shawe-Taylor. Canonical correlation analysis; an overview with application to learning methods. *Neural Computation*, 16(12):2639–2664, 2004.

[6] T-K. Kim, J. Kittler, and R. Cipolla. Learning discriminative canonical correlations for object recognition with image sets. In *European Conference on Computer Vision (ECCV)*, pages 251–262, 2006.

[7] R. Donner, M. Reiter, G. Langs, P. Peloscheck, and H. Bischof. Fast active appearance model search using canonical correlation analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10):1690–1694, October 2006.

[8] M. Reiter, R. Donner, G. Langs, and H. Bischof. 3d and infrared face reconstruction from rgb data using canonical correlation analysis. In *International Conference on Pattern Recognition (ICPR)*, 2006.

[9] T-K. Kim, O. Arandjelovic, and R. Cipolla. Learning over sets using boosted manifold principal angles (bompa). In *British Machine Vision Conference (BMVC)*.

[10] M. Nishiyama, O. Yamaguchi, and K. Fukui. Face recognition with the multiple constrained mutual subspace method. In *International Conference on Audio and Video-Based Person Authentication (AVBPA)*, pages 71–80, 2005.

[11] J. Yang, A. F. Zhang, D. Frangi, and J. Yang. Two dimensional pca: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):131–137, 2004.

[12] J. Ye. Generalized low rank approximations of matrices. *Machine Learning*, 61:167–191, November 2005.

[13] H. Kong, L. Wang, E. K. Teoh, X. Li, J. Wang, and R. Venkateswarlu. Generalized 2d principal component analysis for face image representation and recogniton. *Neural Networks*, 18:585–594, 2005.

[14] J. Ye, R. Janardan, and Q. Li. Two-dimensional linear discriminant analysis. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1569–1576, 2004.

[15] H. Kong, L. Wang, E. K. Teoh, X. Li, J. Wang, and R. Venkateswarlu. A framework of 2d fisher discriminant analysis: Application to face recognition with small number of training samples. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

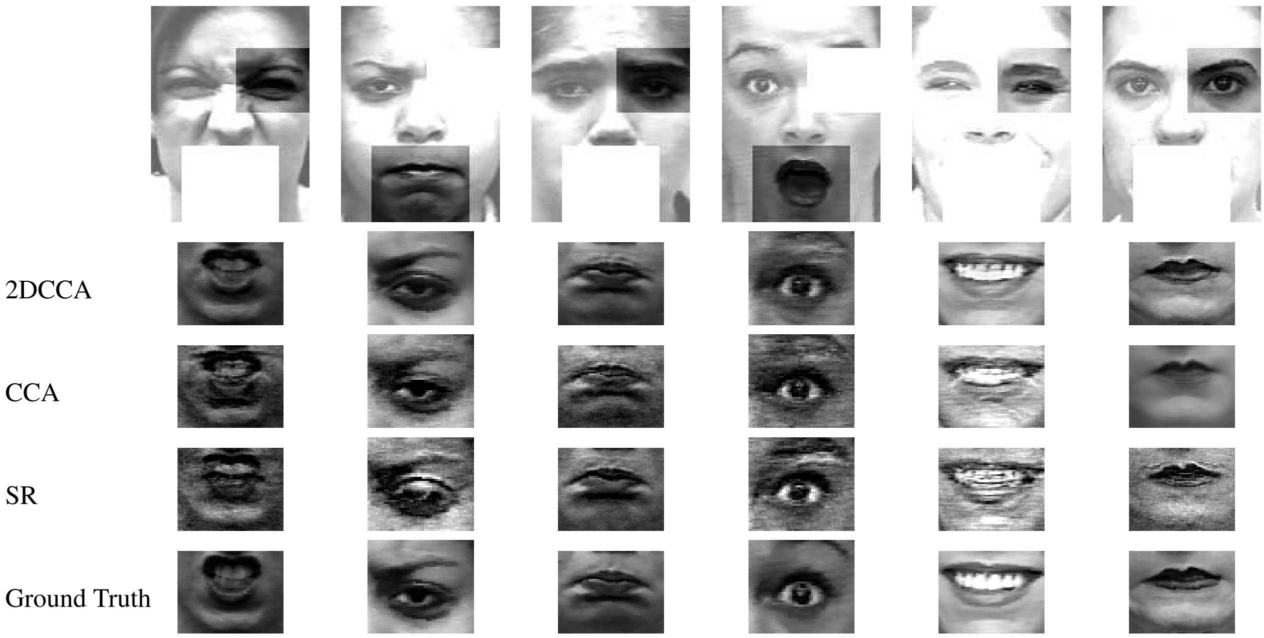

Figure 1: Some examples of facial parts synthesis using 2DCCA, CCA, and SR.