

PREDICTIVE MODELING OF ANATOMIC STRUCTURES USING CANONICAL CORRELATION ANALYSIS

Tianming Liu, Dinggang Shen, Christos Davatzikos

Section of Biomedical Image Analysis, Department of Radiology
University of Pennsylvania, Philadelphia, PA 19104

ABSTRACT

In this paper, we present a method for predictive modeling of anatomic structures using canonical correlation analysis (CCA). Using this technique, certain anatomical structures, such as tumor-distorted structures, can be estimated from others by exploring the correlation between them, which has been determined from a set of training samples. Cortical surfaces and corpus callosum boundaries have been used to demonstrate the performance of the proposed method in predictive modeling. Applications of this method are in estimating brain tissues obscured by tumors and surrounding edema, in detecting abnormal structures, and in formulating alternate forms of statistically-based interpolation and regularization.

1. INTRODUCTION

Predictive modeling, i.e., predicting invisible anatomic structures from the visible structures, has a wide range of medical applications. For example, in neurosurgical planning, it is of interest to register an atlas to the image of a patient with a tumor in order to transfer the information available in the atlas to the patient image. However, reconstructing structures directly from areas severely distorted by tumors or otherwise obscured by edema would prove difficult. Therefore, we attempt to predict the invisible anatomic structures from the rest of normal structures by exploring their correlation. Naturally, application of the predictive modeling to images with tumors also requires that the mass effect is accounted for in the statistical model as in [7].

Predictive modeling may also be applicable in structure abnormality detection and in formulating statistically-based interpolation and regularization schemes. The underlying premise in using predictive modeling to detect structural abnormality is that abnormality can be defined as significant deviation of part of the anatomy from what the remaining normal anatomy predicts for that part. Finally, predictive modeling quantifies the correlation between neighboring or distant anatomic structures. Hence, it could be used as a statistically-based interpolate from discrete data as well as for regularization.

2. ANATOMIC CORRELATION

The predictability between variables greatly depends on their inter-correlation, that is, how well we can estimate the predictant from the predictor depends on how great an inter-correlation exists. Fortunately, there usually exists strong correlation between anatomic structures, which is referred to as anatomic correlation

and is often used as prior information for the segmentation of anatomic structures (c.f. [1]).

Anatomic correlation is defined as follows. Denote an anatomic structure by S , and its vertices by $V = \{v_i \mid v_i \in S\}$. The anatomic correlation between any two vertices v_i and v_j is defined as

$$\text{corr}^x(v_i^x, v_j^x) = \frac{\text{COV}(v_i^x, v_j^x)}{\sigma_i^x \sigma_j^x} \quad (1)$$

where σ_i^x and σ_j^x are standard deviations of the x-coordinate v_i^x and v_j^x , and $\text{COV}(v_i^x, v_j^x)$ is the covariance of these two x-coordinates. The anatomic correlations for y and z coordinates are similarly defined. Figure 1 shows two examples of color-coded anatomic correlation. We can see that there is strong anatomic correlation within both corpus callosum boundaries and cortical surfaces. This anatomic correlation is the source of predictability between anatomic structures. Our goal here is to utilize the anatomic correlation as effectively as possible, in order to perform the most accurate estimation of invisible structures from the visible ones.

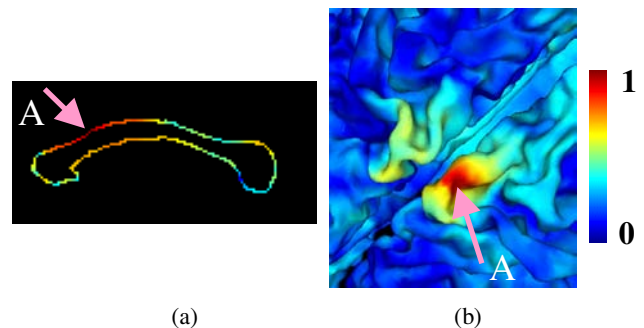


Figure 1. Color-coded anatomic correlations. (a) Color-coded correlation between vertex A and other vertices on the corpus callosum boundary. (b) Color-coded correlation between vertex A and other vertices on the cortical surface.

3. CANONICAL CORRELATION ANALYSIS

Although Hotelling [3] first introduced CCA in 1936, the technique was not brought to application until recently years. Since the late 1980s, CCA has been used extensively in meteorology [5]. The BP-CCA scheme [5] in the Empirical Orthogonal Function (EOF) space was a milestone work because

it avoided the inversion of matrices of non-full rank, which made CCA valuable as an operational forecasting tool. Recently, CCA has been used in the field of biomedical image analysis [6, 8, 10].

CCA explores the correlative structure of two sets of variables. The canonical correlation is the correlation of two canonical (latent) variables, one variable representing a set of independent variables, and the other representing a set of dependent variables. The canonical correlation is optimized such that the linear correlation between the two latent variables is maximized. There may be more than one such linear correlation relating the two sets of variables, with each such correlation representing a different dimension by which the independent set of variables is related to the dependent set. CCA offers an exploratory, dimensionality-reducing technique that is also predictive.

Canonical correlation is a member of the multiple general linear hypothesis family and shares many of the assumptions of linear models, such as linearity of relationships, low multicollinearity, homoscedasticity, multivariate normality, non-singularity in the correlation matrix, a sufficient number of training cases, no outliers, and no error in the independent variables. In this paper, these assumptions are satisfied or are assumed to be satisfied.

The mathematical introduction of CCA is described as follows. Assuming two sets of variables, denoted as

$$\{X = [x_1, x_2, \dots, x_p]^T\} \text{ and } \{Y = [y_1, y_2, \dots, y_q]^T\}$$

CCA extracts the correlated modes between vectors X and Y , by seeking a set of transformation vector pairs, A_i and B_i , which yields the canonical variates u_i and v_i with maximum correlation:

$$u_i = X^T A_i \quad \text{and} \quad v_i = Y^T B_i \quad (2)$$

Here, symbol i denotes the i -th pair of transformation vectors, A_i and B_i , resulting in the i -th pair of variates u_i and v_i . Unlike PCA, which examines a single set of variables to determine modes of maximum variance, CCA examines two sets of variables to determine modes of maximum correlation. According to Eq. (1), the correlation between u_i and v_i is

$$\rho_i = \frac{A_i^T C_{XY} B_i}{\sqrt{A_i^T C_{XX} A_i B_i^T C_{YY} B_i}} \quad (3)$$

where C_{XY} is the cross-covariance matrix of X and Y , C_{XX} and C_{YY} are auto-covariance matrix. To maximize Eq. (3), we obtain the partial derivative of ρ_i with respect to A_i and set the derivative to be zero. So we have

$$C_{XY} B_i = \frac{A_i^T C_{XY} B_i}{A_i^T C_{XX} A_i} C_{XX} A_i \quad (4)$$

Similarly, by setting the partial derivative of ρ_i with respect to B_i to be zero, we have

$$C_{YX} A_i = \frac{B_i^T C_{YX} A_i}{B_i^T C_{YY} B_i} C_{YY} B_i \quad (5)$$

Combining equations (4) and (5), we obtain

$$\begin{cases} C_{XX}^{-1} C_{XY} C_{YY}^{-1} C_{YX} A_i = \rho_i^2 A_i \\ C_{YY}^{-1} C_{YX} C_{XX}^{-1} C_{XY} B_i = \rho_i^2 B_i \end{cases} \quad (6)$$

By solving the eigenvalue problem in Eq. (6), we obtain the decreasingly sorted correlations $\{\rho_1, \rho_2, \dots, \rho_n\}$, and the corresponding transformation vectors, $\bar{A} = [A_1, A_2, \dots, A_n]$ and $\bar{B} = [B_1, B_2, \dots, B_n]$. Also, the corresponding sets of canonical variates can be expressed as $U = [u_1, u_2, \dots, u_n]^T$ and $V = [v_1, v_2, \dots, v_n]^T$.

CCA simplifies the correlation structure among predictor and predictant vectors, as illustrated in Figure 2a. In the original space, for any element x_i of vector X , this element has correlation with any other element in vector X and vector Y . This is also true for any element y_i of Y . But in the canonical space, the elements within U or V are orthogonal. Moreover, any element u_i of vector U is correlated only with the corresponding element v_i of V . Importantly, the correlation between u_i and v_i is maximized in the CCA transformation. Hence, the predictability between u_i and v_i is maximized.

After the CCA transformation, the estimation is performed in the canonical space. A linear regression is used for the estimation of u_i from v_i or the reverse. Figure 2b illustrates the estimation scheme. The estimation of v_1 from u_1 is the most accurate as the correlation between v_1 and u_1 is the maximal. Notably, v_1 is somewhat similar to the first principle component in PCA, and thus is the most important for the reconstruction of a new vector Y' . However, CCA maximizes the predictability of v_1 from u_1 , unlike PCA which is optimal for representation and not for prediction. Similarly, the estimation of v_2 is the second-most accurate and the second-most important for the reconstruction of Y' demonstrating that more important components offer better estimation accuracy in the CCA-based estimation. A flowchart illustrating the estimation of Y' from X' is shown in Figure 2b.

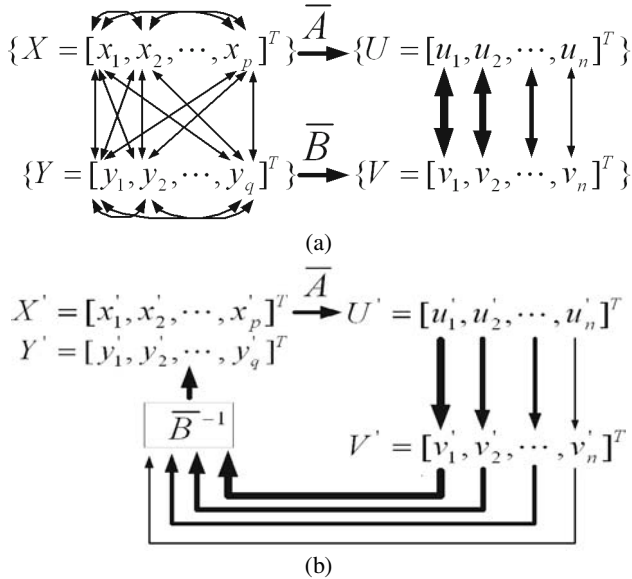


Figure 2. Explanation of the CCA-based estimation scheme. (a) Training stage. The two-way arrow denotes inter-correlation, with the thicker arrow indicating higher correlation. (b) Estimation stage. The thicker arrow denotes a more accurate or more important estimation.

In Figures 2, the CCA is performed directly on the predictant Y and the predictor X . However, two implementation issues arise. Firstly, the CCA scheme requires that the covariance matrices of X and Y be of full rank. However, this does not hold in our application, since there always exist multicollinearities, such as inter-correlations between the variables within predictor or predictant vectors. Secondly, as Y and X have thousands of variables, the computational burden is very heavy.

As mentioned earlier, the BP-CCA [5] was a milestone work because it was performed in the Empirical Orthogonal Function space to avoid the inversion of matrices of non-full rank. In our approach, the EOF space is the principle component eigen space. Therefore, PCA is first performed for Y and X , which are then represented as

$$Y = E_Y \cdot \beta_Y \text{ and } X = E_X \cdot \beta_X \quad (7)$$

where β_Y and β_X correspond to the reconstruction coefficients, and E_Y and E_X are the eigenvector matrices. Then the estimation method based on CCA introduced above is applied to β_Y and β_X . Notably, performing PCA before CCA removes multicollinearities and thus avoids the inversion of non-full rank matrices, and also greatly reduces the computational burden.

4. RESULTS

Two experiments were performed, on cortical surfaces and corpus callosum boundaries, respectively, in order to evaluate the CCA-based estimation method.

In the first experiment, the inner cortical surfaces of 60 normal individuals were registered by a method described in [9]. As the goal of the experiment was to evaluate our estimation method, we masked out part of the anatomy lying within a spherical region, and treated that masked anatomy as the unknown anatomy to be estimated. Then the masked anatomy was used as the ground truth to evaluate the estimation method. We compared the results from the CCA method with those of three other methods, the PCR (Principle Component Regression) [4], MLR (Multiple Linear Regression) [4], and PCA-based [7] methods. The four methods were tested using different numbers of training samples, variously sized anatomic masks, and various positions for those masks. Of the four methods, CCA produces the lowest estimation error, as can be seen in Figures 3, 4 and 5. Figure 6 shows an example of color-coded errors by the CCA-based method for visual evaluation.

In the second experiment, 100 corpus callosum boundaries were manually delineated by experts, and their correspondences were obtained using the constant-speed parameterization method [2]. As in the first experiment, we masked out part of the anatomy lying within a circle region, and treated this region as the unknown anatomy to be estimated. Then the masked anatomy was used as the ground truth to evaluate the estimation method. Again, it can be seen that the CCA-based method performs most accurately, when applied to a varied number of training samples, as shown in Figure 7. Figure 8 visually shows the masked segments and their estimated segments at two different positions.

5. CONCLUSION

A method for predictive modeling of anatomic structures using CCA is presented. Experimental results have shown that it is possible to accurately estimate large areas of invisible anatomic structures from the visible structures. Our future work will use the predictive modeling in support of neurosurgical planning, structure abnormality detection, and statistically-based interpolation and regularization.

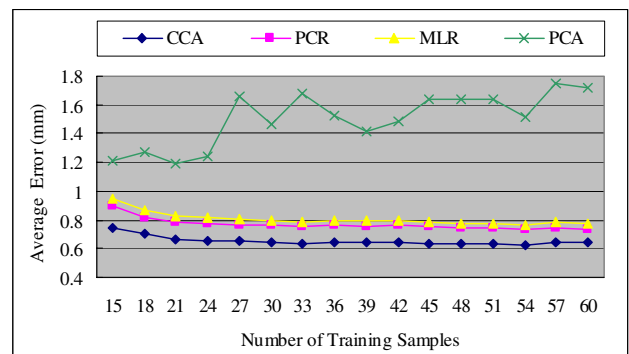


Figure 3. Average estimation errors vary depending on the how many training samples are used. Here, the diameter of the mask, within which the anatomy is estimated, is 30 mm.

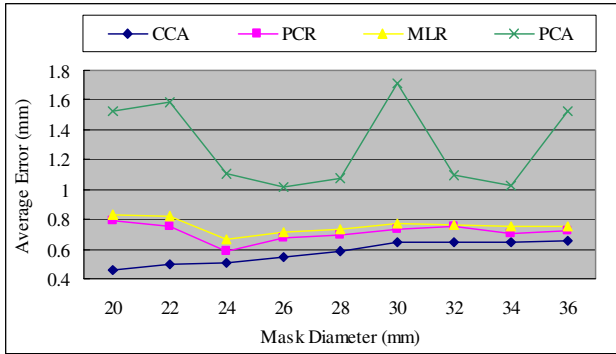


Figure 4. Average estimation errors vary with mask diameter.

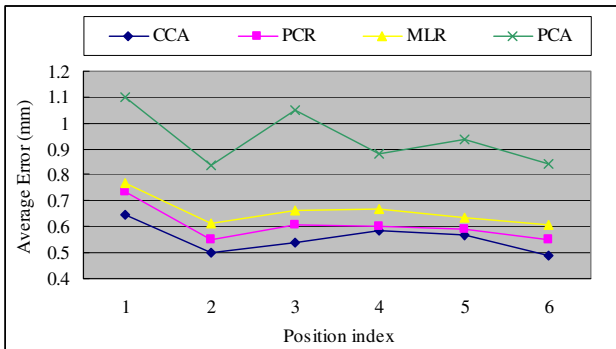


Figure 5. Average estimation errors vary as the mask position is varied. Here, the diameter of the mask is 30 mm.

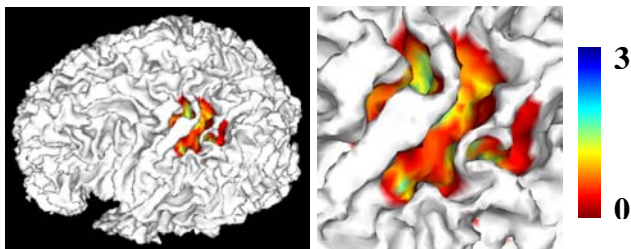


Figure 6. Color-coded estimation errors by our method. The diameter of spherical mask is 40 mm. Average error is 0.52 mm.

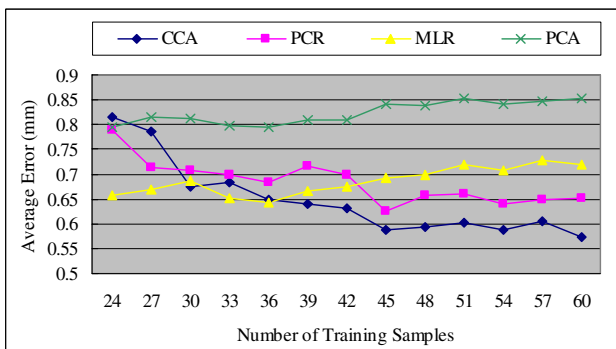


Figure 7. Average estimation errors resulting from each of the four methods vary with the number of corpus callosum samples used. The masked segments can be seen as red curves in Figure 8a.

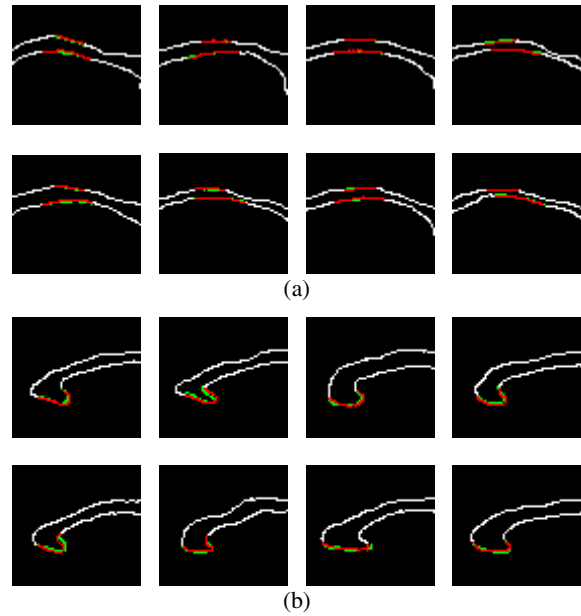


Figure 8. The masked segments (red) and the resulting estimations (green), at two different positions. The white curve is the known part of the structure.

References

1. J. Yang, L. H. Staib, and J. S. Duncan, "Statistical Neighbor Distance Influence in Active Contours," MICCAI, Tokyo, Japan, pp. 588-595, Sept. 25-28, 2002.
2. C. Davatzikos, X. Tao, and D. Shen, "Hierarchical Active Shape Models, Using the Wavelet Transform," IEEE Trans. on Med. Imaging, 22(3):414-423, March, 2003.
3. H. Hotelling, "Relations between Two Sets of Variates," Biometrika, 28, 321-377, 1936.
4. J. O. Rawlings, S. G. Pantula, and D. A. Dickey, "Applied Regression Analysis: a Research Tool," New York: Springer, 1998.
5. T. P. Barnett and R. Preisendorfer, "Origins and Levels of Monthly and Seasonal Forecast Skill for United States Surface Air Temperatures Determined by Canonical Correlation Analysis," Monthly Weather Review, 115, 1825-1850, 1987.
6. O. Friman, J. Carlsson, P. Lundberg, M. Borga, and H. Knutsson, "Detection of Neural Activity in Functional MRI using Canonical Correlation Analysis," Magnetic Resonance in Medicine, Volume 45, Issue 2, 2001.
7. C. Davatzikos, D. Shen, A. Mohamed, and S. Kyriacou, "A Framework for Predictive Modeling of Anatomical Deformations," IEEE Trans. on Med. Imaging, 20 (8):836-843, 2001.
8. M. Ragnehed, O. Friman, P. Lundberg, B. Söderfeldt, and H. Knutsson, "Comparing CCA and SPM," Human Brain Mapping, New York, June 18-22, 2003.
9. T. Liu, D. Shen, and C. Davatzikos, "Deformable Registration of Cortical Structures via Hybrid Volumetric and Surface Warping," MICCAI, Canada, Nov 16-18, 2003.
10. L. Zöllei, L. Panych, E. Grimson, W. Wells, "Exploratory Identification of Cardiac Noise in fMRI Images," MICCAI, Canada, Nov 16-18, 2003.