Dependently Coupled Principal Component Analysis (DC-PCA) for Bivariate Inversion Problems

Navdeep Dahiya,[1] Yifei Fan,[1] Samuel Bignardi,[1] Romeil Sandhu,[2] and Anthony Yezzi[1]

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology; [2]Computer Science Department, Stony Brook University

ABSTRACT

We introduce a new method of leveraging PCA between paired datasets in a dependently coupled manner, which is optimal with respect to approximation error during training.

- PLSR/CCA [1,2] maximize $cov/corr$ b/w paired data, symmetrically without differentiating b/w observable and unobservable data.
- DC-PCA generates dependently coupled paired basis by relaxing orthogonality constraints in decomposing unobservable component.
- DC-PCA better suited for inversion problems.

OBJECTIVES

- In typical inversion problems, full inversion of variable X with limited observability is impractical.
- Obtain estimate of X based on low dimensional inversion for variable X.
- Obtain a low dimensional estimate of hidden/unobservable variable (Y) based on the low dimensional inversion of the coupled variable (X).

PROPOSED DC-PCA

MxN matrix $X = [\mathbf{x}_1\ \mathbf{x}_2 \cdots \mathbf{x}_N]$ with N data measurements, and KxN matrix $Y = [\mathbf{y}_1\ \mathbf{y}_2 \cdots \mathbf{y}_N]$ with N individually paired measurements.

Standard PCA on X minimizes MSE: $\varepsilon_X(A, U) = \frac{1}{N}\sum_{n=1}^{N} \|\mathbf{x}_n - U\boldsymbol{a}_n\|^2$ where U is an orthonormal basis, and a_n are expansion coefficients for closest approximations to measurements x_n in this basis U.

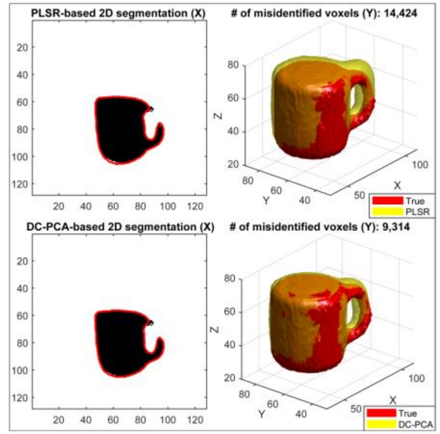
If we impose basis U together with optimal coefficients, then we may seek paired basis V (not necessarily orthogonal), that minimizes MSE:

$$\varepsilon_Y(A, V) = \frac{1}{N}\sum_{n=1}^{N} \|\mathbf{y}_n - V\boldsymbol{a}_n^*\|^2$$

Using $\boldsymbol{a}_n^*(\mathrm{U}) = U^T \mathrm{x}_n$ above, taking derivative wrt V and setting to zero yields:
$V = YX^T U \Lambda_X$ where Λ_X is diagonal matrix of largest eigenvalues of XX^T.

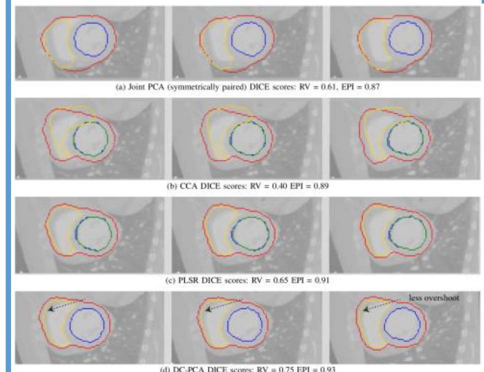
Given expansion coefficients which estimate *observable* X, we may obtain optimal prediction for *unobservable* Y, using same set of coefficients and the paired basis V.

Relationship to PLS/PLSR



- X represents 2D cross-section shape and Y a paired 3D teacup shape.
- Low dimensional shape inversion is applied to ideal noiseless silhouette using both PCA and PLSR basis for X.
- Both methods extract similar shape from raw image data (left).
- However, the estimated 3D surface (Y estimate) from PLSR exhibits higher mismatch against true shape compared to DC-PCA (right).
- This is because the X estimate that would have produced superior PLSR Y estimate is not captured during inversion process.

PRACTICAL APPLICATION



- Segment Left Ventricle (LV), Right Ventricle (RV), and Epicardium (EPI) from cardiac CT imagery using 3D shape models built from manual training segmentations [3].
- LV much easier to segment than RV & EPI (Treat LV as observable X and RV/EPI as coupled unobservables Y).
- Estimate RV (yellow) & EPI (red) using DC-PCA based on fitted LV (blue) shape coefficients.
- DC-PCA does better in estimating unobservable anatomies (less overshoot in RV compared to PLSR/CCA/Joint PCA).

REFERENCES

[1] J. A. Wegelin et al., "A survey of partial least squares (pls) methods, with emphasis on the two-block case," *University of Washington, Tech. Rep,* 2000.

[2] H.HOTELLING,"Relations between two sets of variates." *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.

[3] N. Dahiya, A. Yezzi, M. Piccinelli, and E. Garcia, "Integrated 3D anatomical model for automatic segmentation in cardiact CT imagery," *Computer Methods in Biomechanics & Biomedical Imagery*. 2019.