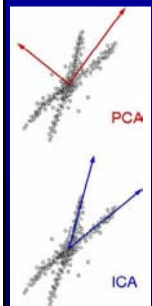
Independent Component Analysis



- *PCA finds the directions that uncorellate*
- *ICA / Blind Source Separation:*
  - Observed data is modeled as a linear combination of independent sources
    - Cocktail Problem: A sound recording at a party is the result of multiple individuals speaking (independent sources)
- *ICA finds the directions of maximum independence*

Computing Independent Components

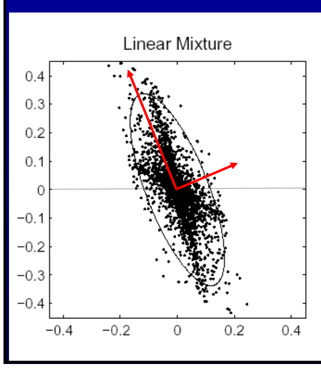
- By maximization of nongaussianity: kurtosis
- By maximum likelihood estimation
- By minimization of mutual information
- By tensorial methods
- By nonlinear decorrelation and nonlinear PCA
- By methods using time structure

• Hyvärinen A, Karhunen J, Oja E. "Independent component analysis", John Wiley & Sons, Inc., New York, 2001, p. 481
• http://www.cis.hut.fi/projects/ica/fastica/

Computing IC's using Non-Gausianity

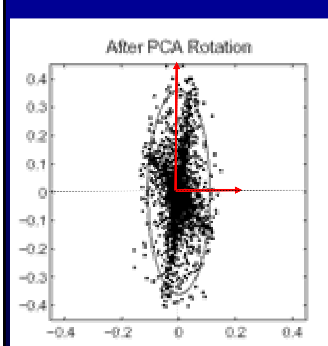
- a measure of non-gaussianity: kurtosis
  - kurt(y) = E{y4} – 3(E{y2})2 = E{y4} – 3
- for unit-variance data
  - kurt(y) = 0 for gaussian data
  - kurt(y) < 0 for subgaussian data
  - kurt(y) > 0 for supergaussian data
- kurtosis is measured along each possible projection direction over the data
  - a maximum corresponds to one of the IC's
  - other IC's are found from the orthogonal directions with an iterative algorithm
  - rotation matrix R has now been solved

Geometric View of ICA



$$\mathbf{D} = \mathbf{USV}^T$$

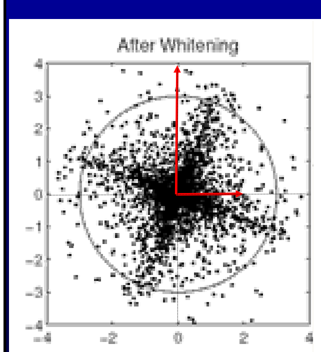
Geometric View of ICA



$$\mathbf{D} = \mathbf{USV}^T$$

$$\mathbf{D}' = \mathbf{U}^T\mathbf{D}$$

Geometric View of ICA

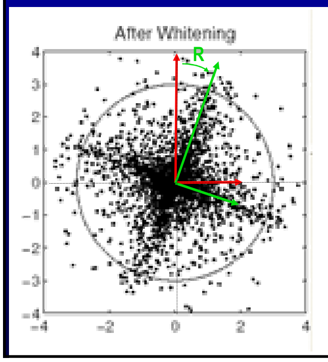


$$\mathbf{D} = \mathbf{USV}^T$$

$$\mathbf{D}' = \mathbf{U}^T\mathbf{D}$$

$$\mathbf{D}'' = \mathbf{S}^{-\frac{1}{2}}\mathbf{U}^T\mathbf{D}$$

Geometric View of ICA



$$\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

$$\mathbf{D}' = \mathbf{U}^T\mathbf{D}$$

$$\mathbf{D}'' = \mathbf{S}^{-\frac{1}{2}}\mathbf{U}^T\mathbf{D}$$

$$\mathbf{D}''' = \mathbf{R}\mathbf{S}^{-\frac{1}{2}}\mathbf{U}^T\mathbf{D}$$

$$\mathbf{D} = \mathbf{U}\underbrace{\mathbf{S}^{\frac{1}{2}}\mathbf{R}^T}\underbrace{\mathbf{R}\mathbf{S}^{-\frac{1}{2}}}\mathbf{S}\mathbf{V}^T$$

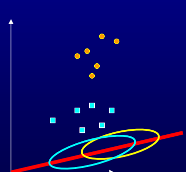
$$\mathbf{D} = \underbrace{\mathbf{U}\mathbf{W}^{-1}}\mathbf{W}\mathbf{S}\mathbf{V}^T$$

Independent Components

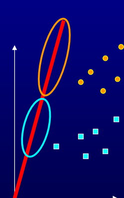
Fisher Linear Discriminant: FisherFaces

Fisher's Linear Discriminant

- Objective: Find a projection which separates data clusters



Poor separation



Good separation

FLD: Data Scatter

- Within class scatter matrix

$$\mathbf{S}_W = \sum_{c=1}^{C} \sum_{\mathbf{i}_n \in D_c} (\mathbf{i}_n - \mu_c)(\mathbf{i}_n - \mu_c)^T$$

- Between class scatter matrix

$$\mathbf{S}_B = \sum_{c=1}^{C} |D_c| (\boldsymbol{\mu}_c - \boldsymbol{\mu})(\boldsymbol{\mu}_c - \boldsymbol{\mu})^T$$

- Total scatter matrix

$$\mathbf{S}_T = \mathbf{S}_W + \mathbf{S}_B$$

Fisher Linear Discriminant

- The basis matrix B is chosen in order to maximize ratio of the determinant between class scatter matrix of the projected samples to the determinant within class scatter matrix of the projected samples

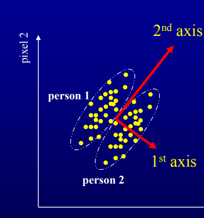
$$\mathbf{B} = \arg\max_{\mathbf{B}} \frac{|\mathbf{B}^T\mathbf{S}_{btw}\mathbf{B}|}{|\mathbf{B}^T\mathbf{S}_{in}\mathbf{B}|}$$

- B is the set of generalized eigenvectors of S_{Btw} and S_{Win} corresponding with a set of decreasing eigenvalues

$$\mathbf{S}_{btw}\mathbf{B} = \mathbf{S}_{within}\Lambda\mathbf{B}$$

Fisher Linear Discriminant

- Consider a set of images of 2 people under fixed viewpoint & N lighting condition



- Each image is represented by one coefficient vector
- Each person is displayed in N images and therefore has N coefficient vectors

© 2002 by M. Alex O. Vasilescu