Department of Computer Engineering
IC@Network Lab

Cognitive States Detection in fMRI Data Analysis using incremental PCA

Hoang Trong Minh Tuan, Yonggwan Won*, Hyung-Jeong Yang

International Conference on Computational Science and its Applications (ICCSA 2007)

Agenda

- Chapter 1: Introduction
- Chapter 2: Data Description
- Chapter 3: Previous works
- Chapter 4: iPCA
- Chapter 5: Experimental results

functional Magnetic Resonance Imaging

Chapter 1

Introduction

- Application
  - Investigating internal activities in human brain.
  - Disease detection (Alzheimer, Parkinson)
- How
  - Measure changes in blood oxygen level dependent (BOLD) response using fMRI scanner
- Data result
  - Three-dimensional images showing changes in BOLD via time

functional Magnetic Resonance Imaging

Chapter 1

Introduction

- Research interest:
  - Obtain a comprehensive understanding of human brain activity – elucidating the relationship between structure and function.
  - Identify the cognitive states from fMRI images
- Data characteristics
  - Sparse: (dozens of data) hard and costly to collect data
  - High-dimensions: up to hundred thousands of voxels
  - Noises: nontask-related brain tissue activations + head motion + machine artifact, etc.

Data description

- Input-output mapping

$$f: fMRI - sequence - images(t_1, t_2) \rightarrow Cognitive - States$$

  - fMRI-sequence-images(t_1,t_2) : a sequence of time-series fMRI images collected during interval [t_1, t_2]
  - Cognitive-States: a set of cognitive states to be detected
- Information
  - Picture vs. Sentence
  - t_2-t_1 = 8s

Chapter 2 Data description

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Data description

- Sentence
  - Positive: “It is true that the star is below the plus”
  - Negative: “It is not true that the star is above the plus”
- Pictures
  - Symbols: +, *, - ,…

Table 1. Dataset information

Subject ID	04799	04820	04847	05675	05680	05710
# features	79184	80240	75168	82160	80992	74144

Chapter 2 Data description

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Data description

Chapter 2

Data description

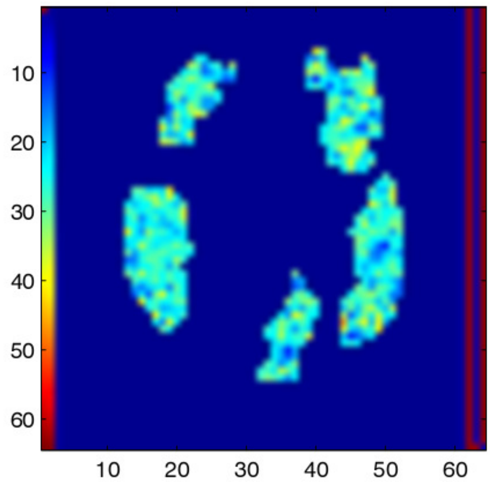

Figure 1. fMRI image of the subject ID 05710 when reading a sentence 'It is not true that the star is above the plus.' in a trial at slice z=6 of 8-th snapshot.

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Previous works

Chapter 3

Previous works

- Unsupervised learning method: clustering
- Independent Component Analysis
- Select *n* most discriminating voxels
- Select *n* most active voxels
- Select *n* most active voxels per Region of Interests
- PCA (on powerful computers)

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Incremental Principal Component Analysis

Chapter 4

iPCA

- Observation = *n*-dimensional vector **x** (feature vector)
- Feature vector = point in *vector space* $\mathbb{R}^n$.
- Feature vector = linear combination of a basis $\mathcal{V}$

$$\mathcal{V} = \{\vec{\mathbf{v}_1}, \vec{\mathbf{v}_2}, ..., \vec{\mathbf{v}_n}\}$$

$$\mathbf{x} = c_1\vec{\mathbf{v}_1} + c_2\vec{\mathbf{v}_2} + ... + c_n\vec{\mathbf{v}_n}$$

- Normally, standard basis E is used

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Principal Component Analysis

Chapter 4

iPCA

- Problem
  - Standard basis → unknown ideal basis
  - Keep most *p* significant principal components of the new eigenspace
- Incremental PCA:
  - One-by-one update the $v_{i,j}$ to minimize the reconstruction error

$$\sum_k \|(\tilde{\mathbf{x}}_k - \mathbf{x}_k)\|$$

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Incremental Principal Component Analysis

Chapter 4

iPCA

- Assume
  - $\vec{\mathbf{v}_i} = [v_{i,1}, ..., v_{i,n}]^T \ (i = 1...p)$ eigenspace
  - $\mathbf{a}_k$ the projection of $\mathbf{x}_k$ onto the *p*-dimensional eigenspace (k=1…m)

$$a_{k,i} = v_{i,1}.x_{k,1} + v_{i,2}.x_{k,2} + ... + v_{i,n}x_{k,n}$$

with $i = 1...p$.

  - $\tilde{\mathbf{x}}_k = [\tilde{x}_{k,1}, \tilde{x}_{k,2}, ..., \tilde{x}_{k,n}]$ the reconstructed point

$$\tilde{x}_{k,j} = v_{1,j}.a_{k,1} + v_{2,j}.a_{k,2} + ... + v_{p,j}.a_{k,p}$$

with $j = 1...n$.

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Incremental Principal Component Analysis

Chapter 4

iPCA

- Basic idea: for each new sample $\mathbf{x}_{k+1}$

1. Computer $\acute{\mathbf{a}}_{k+1}$ by projecting $\mathbf{x}_{k+1}$ on to the current principal components $\vec{\mathbf{v}_i}$ $(i = 1...p)$
2. Estimate the residue vector (reconstruction error) $e_i = \|\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}\|$ $(i = 1...p)$
3. Update the estimates of $\vec{\mathbf{v}_i}$, $(i = 1...p)$, and output the actual projected point for sample $\mathbf{x}_{k+1}$: $\mathbf{a}_{k+1}$

- Magnitudes of updates
  - is proportional to the residual vector
  - is inversely proportional to the past data captured

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Incremental Principal Component Analysis

Chapter 4

iPCA

Algorithms

Step 1: Initialize the p principal components $\vec{\mathbf{v}_i}$ to unit vectors, i.e. $\vec{\mathbf{v}_1} = [100...0]^T, \vec{\mathbf{v}_2} = [010...0]^T, \vec{\mathbf{v}_3} = [001...0]^T$, etc., and d_i to small positive values.

Step 2: For each samples $\mathbf{x}_{k+1}$, assign $\mathbf{u}_1 = \mathbf{x}_{k+1}$, and start updating $\mathcal{V}$

Step 2.1: for each principal components $\vec{\mathbf{v}_i}$ $(i = 1...p)$, do

- ▶ find the component $a_{k+1,i}$ by projecting $\mathbf{x}_{k+1}$ onto $\vec{\mathbf{v}_i}$
- ▶ estimate the residue vector (error) $\mathbf{e} = \mathbf{u}_i - a_{k+1,i}.\vec{\mathbf{v}_i}$
- ▶ calculate the past data captured by $\vec{\mathbf{v}_i}$: $d_i \leftarrow \lambda d_i + a^2_{k+1,i}$
- ▶ update the principal component: $\vec{\mathbf{v}_i} \leftarrow \vec{\mathbf{v}_i} + \frac{a_{k+1,i}.\mathbf{e}_i}{d_i}$
- ▶ repeat with the remainder of $\mathbf{u}_{i+1}$: $\mathbf{u}_{i+1} = \mathbf{u}_i - a_{k+1,i}.\vec{\mathbf{v}_i}$

Step2.2: do re-orthonormalization to maintain orthonormality of the system.

Step 3: Finally, project the whole data to the new eigenspace model with p principal components

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Experiment process

Chapter 5

Experiments

- Pre-process:
  - Remove noise
  - Collect task-related images only
  - Classify data to 'Picture vs. Sentence' study only
- Gaussian Naïve Bayes classifier
  - Classification errors are computed using k-fold cross-validation

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Picture-Sentence study

Chapter 5

Experiments

- Apply for full feature set and compare with *p*=250 principal components

Table 2. Single subject cognitive state detector using GNB for Picture vs. Sentence study using iPCA

Feature selection	Average Performance	Subject					
		04799	04820	04847	05675	05680	05710
All features ($\sim$ 80,000)	63.75%	53.75%	57.50%	75.00%	58.75%	67.50%	70.00%
iPCA(250)	83.75%	80.00%	80.00%	90.00%	88.75%	78.75%	85.00%

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Picture-Sentence study

Chapter 4

Experiments

- Compare with previous methods

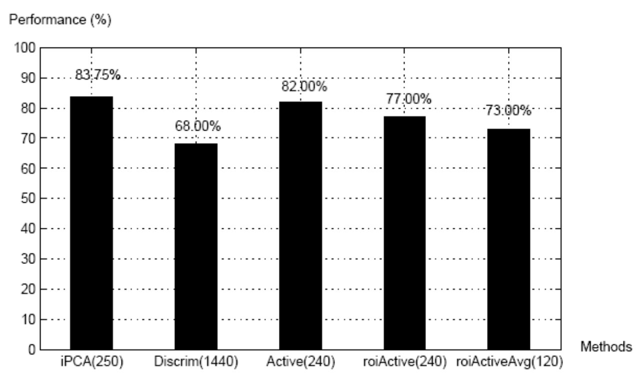

Figure 3. Comparison between iPCA and other feature selection methods using GNB classifier for 'Picture vs. Sentence' study

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Tasks in order study

- Consider trials in Picture - Sentence order

Table 3. Experimental results for trials with Picture-Sentence sequence study using iPCA

Feature selection	Average Performance	Subject					
		04799	04820	04847	05675	05680	05710
iPCA(250)	87.08%	97.50%	87.50%	80.00%	80.00%	87.50%	90.00%

Chapter 4

Experiments

- Consider trials in Sentence – Picture order

Table 4. Experimental results for trials with Sentence-Picture sequence study using iPCA

Feature selection	Average Performance	Subject					
		04799	04820	04847	05675	05680	05710
iPCA(250)	92.91%	100.00%	95.00%	85.00%	92.50%	92.50%	92.50%

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Utilize regions of Interests

- In brains, there are regions which active the most for a particular task → Regions of Interest

Chapter 4

Experiments

Table 5. Comparison between iPCA and Domain-Expert-Depending-Selection (DEDS) method

Feature Selection method	Average Performance	
	Picture-Sentence sequence	Sentence-Picture sequence
iPCA(250)	0.87±0.06	0.93±0.05
DEDS(8,470∼11,136)	0.80±0.034	0.90±0.26

Table 6. Single-subject cognitive state detector using GNB for Affirmative vs. Negative Sentence study using iPCA

Feature selection	Average Performance	Subject					
		04799	04820	04847	05675	05680	05710
All features (∼ 80,000)	58.33%	72.50%	62.50%	47.50%	55.00%	55.00%	57.50%
iPCA(250)	80.41%	75.00%	82.50%	77.50%	92.50%	80.00%	75.00%

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab

Conclusion

Chapter 5

Conclusion

- Dimension reduction or knowledge extraction from fMRI data is considered as a challenged problem.
- Only need to keep the first *p* eigenstructures → algorithms can be applied to personal computer.
- The temporal and spatial data fMRI's feature vectors can be reduced.
- The results proved the new eigenspace can capture enough information.

Thank you!

Chonnam National University - KOREA
Dept. of Computer Engineering
IC@Network Lab