A parcellation scheme based on von Mises-Fisher distributions and Markov random fields for segmenting brain regions using resting-state fMRI

Shikhar Agrawal and Mayank Jain

Supervised by: Prof. Suyash P. Awate

Indian Institute of Technology Bombay

29 April, 2022

Table of Contents

Table of Contents

VMF-MRF Brain Parcellation using resting-state fMRI

- Need for investigating the functional subdivisions of the human brain

VMF-MRF Brain Parcellation using resting-state fMRI

- Need for investigating the functional subdivisions of the human brain
- A novel spatio-temporal probabilistic parcellation scheme that overcomes major weaknesses of existing approaches by-

VMF-MRF Brain Parcellation using resting-state fMRI

- Need for investigating the functional subdivisions of the human brain
- A novel spatio-temporal probabilistic parcellation scheme that overcomes major weaknesses of existing approaches by-
- Modeling the **fMRI time series of a voxel as a von Mises-Fisher distribution**, which is widely used for clustering high dimensional data

VMF-MRF Brain Parcellation using resting-state fMRI

- Need for investigating the functional subdivisions of the human brain
- A novel spatio-temporal probabilistic parcellation scheme that overcomes major weaknesses of existing approaches by-
- Modeling the **fMRI time series of a voxel as a von Mises-Fisher distribution**, which is widely used for clustering high dimensional data
- Modeling the **latent cluster labels as a Markov random field**, which provides spatial regularization on the cluster labels by penalizing neighboring voxels having different cluster labels

VMF-MRF Brain Parcellation using resting-state fMRI

- Need for investigating the functional subdivisions of the human brain
- A novel spatio-temporal probabilistic parcellation scheme that overcomes major weaknesses of existing approaches by-
- Modeling the **fMRI time series of a voxel as a von Mises-Fisher distribution**, which is widely used for clustering high dimensional data
- Modeling the **latent cluster labels as a Markov random field**, which provides spatial regularization on the cluster labels by penalizing neighboring voxels having different cluster labels
- Introducing a prior on the number of labels, which helps in **uncovering the number of clusters automatically from the data**

VMF-MRF Brain Parcellation using resting-state fMRI

- Need for investigating the functional subdivisions of the human brain
- A novel spatio-temporal probabilistic parcellation scheme that overcomes major weaknesses of existing approaches by-
- Modeling the **fMRI time series of a voxel as a von Mises-Fisher distribution**, which is widely used for clustering high dimensional data
- Modeling the **latent cluster labels as a Markov random field**, which provides spatial regularization on the cluster labels by penalizing neighboring voxels having different cluster labels
- Introducing a prior on the number of labels, which helps in **uncovering the number of clusters automatically from the data**
- Modeling **spatial correlations in rs-fMRI data** to produce a parcellation even for non-contiguous regions of human brain

Existing Approaches and Shortcomings

- The **Positron Emission Tomography** based **invasive Cytoarchitectonic Mappings** in postmortem brains is a solution

Existing Approaches and Shortcomings

- The **Positron Emission Tomography** based **invasive Cytoarchitectonic Mappings** in postmortem brains is a solution
- Anatomical parcellations using **cellular organization, gyral folding patterns, or neurotransmitter profiles** are examples of above

Existing Approaches and Shortcomings

- The **Positron Emission Tomography** based **invasive Cytoarchitectonic Mappings** in postmortem brains is a solution
- Anatomical parcellations using **cellular organization, gyral folding patterns, or neurotransmitter profiles** are examples of above
- Anatomical parcellations in these atlases are based on postmortem brains of the elderly, and are **less useful** for examining individual differences in the functional organization of the normal healthy adult brain

Existing Approaches and Shortcomings

- The **Positron Emission Tomography** based **invasive Cytoarchitectonic Mappings** in postmortem brains is a solution
- Anatomical parcellations using **cellular organization, gyral folding patterns, or neurotransmitter profiles** are examples of above
- Anatomical parcellations in these atlases are based on postmortem brains of the elderly, and are **less useful** for examining individual differences in the functional organization of the normal healthy adult brain
- Parcellation of the cortex into subnetworks based on resting-state fMRI data opens up the possibility of developing **novel functional atlases based entirely on cortical function**

More Approaches, More shortcomings

- Currently used methods include **hierarchical clustering, Gaussian mixture modeling , graph theory based spectral clustering, computer vision based edge detection** and the commonly used **K-means based clustering methods**

More Approaches, More shortcomings

- Currently used methods include **hierarchical clustering, Gaussian mixture modeling , graph theory based spectral clustering, computer vision based edge detection** and the commonly used **K-means based clustering methods**
- With the exception of edge detection methods all of these methods **require the specification of number of clusters**

More Approaches, More shortcomings

- Currently used methods include **hierarchical clustering, Gaussian mixture modeling , graph theory based spectral clustering, computer vision based edge detection** and the commonly used **K-means based clustering methods**
- With the exception of edge detection methods all of these methods **require the specification of number of clusters**
- This is a problem because the choice of number of clusters **can significantly influence functional modules uncovered by each method**

More Approaches, More shortcomings

- Currently used methods include **hierarchical clustering, Gaussian mixture modeling , graph theory based spectral clustering, computer vision based edge detection** and the commonly used **K-means based clustering methods**
- With the exception of edge detection methods all of these methods **require the specification of number of clusters**
- This is a problem because the choice of number of clusters **can significantly influence functional modules uncovered by each method**
- A second major limitation of current methods used to parcellate the brain is that they **do not exploit spatial correlations** that are inherent in rs-fMRI data

Table of Contents

Von Mises-Fisher Distribution

Probability Density Function

The probability density function of the von Mises–Fisher distribution for the random p-dimensional unit vector **x** is given by $f_p(x; \mu, \kappa) = C_p(\kappa)e^{\kappa\mu^\top x}$, where $\kappa \geq 0$, $\|\mu\| = 1$ and the normalization constant $C_p(\kappa) = \frac{\kappa^{\frac{p}{2}-1}}{(2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(\kappa)}$, where I_ν denotes the modified Bessel function of the first kind at order ν.

Von Mises-Fisher Distribution

Probability Density Function

The probability density function of the von Mises–Fisher distribution for the random p-dimensional unit vector **x** is given by $f_p(x;\mu,\kappa) = C_p(\kappa)e^{\kappa\mu^\top x}$, where $\kappa \geq 0$, $\|\mu\| = 1$ and the normalization constant $C_p(\kappa) = \frac{\kappa^{\frac{p}{2}-1}}{(2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(\kappa)}$, where I_ν denotes the modified Bessel function of the first kind at order ν.

Relation to Isotropic Covariance MVG

Starting from a normal distribution, with isotropic covariance, $\kappa^{-1}I$, and a mean μ of length $r > 0$ that has the density $G_p(x;\mu,\kappa) = (\sqrt{\frac{\kappa}{2\pi}})^p e^{-\kappa\frac{(x-\mu)^\top(x-\mu)}{2}}$ the Von Mises-Fisher distribution $f_p(x; r^{-1}\mu, r\kappa)$ is obtained by conditioning on $\|x\| = 1$.

Von Mises-Fisher Distribution

Probability Density Function

The probability density function of the von Mises–Fisher distribution for the random p-dimensional unit vector **x** is given by $f_p(x; \mu, \kappa) = C_p(\kappa) e^{\kappa \mu^\top x}$, where $\kappa \geq 0$, $\|\mu\| = 1$ and the normalization constant $C_p(\kappa) = \frac{\kappa^{\frac{p}{2}-1}}{(2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(\kappa)}$, where I_ν denotes the modified Bessel function of the first kind at order ν.

Relation to Isotropic Covariance MVG

Starting from a normal distribution, with isotropic covariance, $\kappa^{-1} I$, and a mean μ of length $r > 0$ that has the density $G_p(x; \mu, \kappa) = (\sqrt{\frac{\kappa}{2\pi}})^p e^{-\kappa \frac{(x-\mu)^\top (x-\mu)}{2}}$ the Von Mises-Fisher distribution $f_p(x; r^{-1}\mu, r\kappa)$ is obtained by conditioning on $\|x\| = 1$.

- μ is the mean parameter and κ is the concentration parameter.

Algorithm

- Given the model parameters, cluster labels are estimated by the **state of the art graph cut method based on an α expansion** scheme

Algorithm

- Given the model parameters, cluster labels are estimated by the **state of the art graph cut method based on an** α **expansion** scheme
- Given the cluster labels for voxels, the **model parameters are estimated by maximizing the log-likelihood**

Algorithm

- Given the model parameters, cluster labels are estimated by the **state of the art graph cut method based on an α expansion** scheme
- Given the cluster labels for voxels, the **model parameters are estimated by maximizing the log-likelihood**
- We will assume that the **parameter for both priors** i.e. spatial smoothness prior and labels prior are known and hence **are manually tuned**(We may relax this assumption later!)

Algorithm

- Given the model parameters, cluster labels are estimated by the **state of the art graph cut method based on an α expansion** scheme
- Given the cluster labels for voxels, the **model parameters are estimated by maximizing the log-likelihood**
- We will assume that the **parameter for both priors** i.e. spatial smoothness prior and labels prior are known and hence **are manually tuned**(We may relax this assumption later!)
- Finding global optimal solutions for the cost functions(**Energy Functions**) that arise from such **discontinuity preserving constraints is NP-hard**, hence we will use the α expansion graph cut minimisation algorithm for **finding approximately optimal cluster labels**

Algorithm

- Given the model parameters, cluster labels are estimated by the **state of the art graph cut method based on an α expansion** scheme
- Given the cluster labels for voxels, the **model parameters are estimated by maximizing the log-likelihood**
- We will assume that the **parameter for both priors** i.e. spatial smoothness prior and labels prior are known and hence **are manually tuned**(We may relax this assumption later!)
- Finding global optimal solutions for the cost functions(**Energy Functions**) that arise from such **discontinuity preserving constraints is NP-hard**, hence we will use the α expansion graph cut minimisation algorithm for **finding approximately optimal cluster labels**
- We will iterate over these steps until the energy function converges.

Notation for this problem

- Let $Y = \{y_i\}_{i=1}^{M}$ be the observed voxel time series, where M is the number of voxels in the region of interest

Notation for this problem

- Let $Y = \{y_i\}_{i=1}^{M}$ be the observed voxel time series, where M is the number of voxels in the region of interest
- Each voxel time series y_i consists of T fMRI observations. We normalize each voxel's time series y_i such that $\|y_i\| = 1$

Notation for this problem

- Let $Y = \{y_i\}_{i=1}^{M}$ be the observed voxel time series, where M is the number of voxels in the region of interest
- Each voxel time series y_i consists of T fMRI observations. We normalize each voxel's time series y_i such that $\|y_i\| = 1$
- Let $X = \{X_i\}_{i=1}^{M}$ be the unknown labels of each voxel where each X_i can take a value in the discrete label set $\mathcal{L} = \{1, 2, 3, ..., L\}$

Notation for this problem

- Let $Y = \{y_i\}_{i=1}^{M}$ be the observed voxel time series, where M is the number of voxels in the region of interest
- Each voxel time series y_i consists of T fMRI observations. We normalize each voxel's time series y_i such that $\|y_i\| = 1$
- Let $X = \{X_i\}_{i=1}^{M}$ be the unknown labels of each voxel where each X_i can take a value in the discrete label set $\mathcal{L} = \{1, 2, 3, ..., L\}$
- Let each T dimensional normalized voxel time series y_i given its cluster label $X_i = l$ follow T-variate VMF distribution given by $p(y_i | X_i = l, \mu_l, \kappa_l) = C(\kappa_l) e^{\kappa_l \mu^\top y_i}$

Notation for this problem

- Let $Y = \{y_i\}_{i=1}^{M}$ be the observed voxel time series, where M is the number of voxels in the region of interest
- Each voxel time series y_i consists of T fMRI observations. We normalize each voxel's time series y_i such that $\|y_i\| = 1$
- Let $X = \{X_i\}_{i=1}^{M}$ be the unknown labels of each voxel where each X_i can take a value in the discrete label set $\mathcal{L} = \{1, 2, 3, ..., L\}$
- Let each T dimensional normalized voxel time series y_i given its cluster label $X_i = l$ follow T-variate VMF distribution given by $p(y_i|X_i = l, \mu_l, \kappa_l) = C(\kappa_l)e^{\kappa_l \mu^\top y_i}$
- We model the prior distribution of unknown cluster labels X as discrete MRF given by $P(X) = \frac{1}{Z}e^{-(U_s(X)+U_l(X))}$, where $U_s(X)$ is the energy function which imposes spatial regularization and $U_l(X)$ is the label cost

Notation for this problem

- $U_s(X) = \sum_{i=1}^{M} \sum_{j \in N_i} V(X_i, X_j)$ where the potts potential function is defined as $V(X_i, X_j) = \beta_s(1 - \delta(X_i - X_j))$, where $\delta(z) = 1$ if $z = 0$ and $\delta(z) = 0$ otherwise

Notation for this problem

- $U_s(X) = \sum_{i=1}^{M} \sum_{j \in N_i} V(X_i, X_j)$ where the potts potential function is defined as $V(X_i, X_j) = \beta_s(1 - \delta(X_i - X_j))$, where $\delta(z) = 1$ if $z = 0$ and $\delta(z) = 0$ otherwise
- The label cost $U_l(X) = \beta_l \sum_{i=1}^{L} \delta_p(X)$, where $\delta_p(X) = 1$, if $\exists$ i such that $X_i = p$ and 0 otherwise

Notation for this problem

- $U_s(X) = \sum_{i=1}^{M} \sum_{j \in N_i} V(X_i, X_j)$ where the potts potential function is defined as $V(X_i, X_j) = \beta_s(1 - \delta(X_i - X_j))$, where $\delta(z) = 1$ if $z = 0$ and $\delta(z) = 0$ otherwise
- The label cost $U_l(X) = \beta_l \sum_{i=1}^{L} \delta_p(X)$, where $\delta_p(X) = 1$, if $\exists$ i such that $X_i = p$ and 0 otherwise
- The posterior distribution of X given the data Y and model parameters $\Theta = [\{\mu_l, \kappa_l\}_{l=1}^{L}]$ is given by $P(X|Y,\Theta) \propto \prod_{i=1}^{M} p(y_i|x_i)P(X)$

Notation for this problem

- $U_s(X) = \sum_{i=1}^{M} \sum_{j \in N_i} V(X_i, X_j)$ where the potts potential function is defined as $V(X_i, X_j) = \beta_s(1 - \delta(X_i - X_j))$, where $\delta(z) = 1$ if $z = 0$ and $\delta(z) = 0$ otherwise
- The label cost $U_l(X) = \beta_l \sum_{i=1}^{L} \delta_p(X)$, where $\delta_p(X) = 1$, if $\exists$ i such that $X_i = p$ and 0 otherwise
- The posterior distribution of X given the data Y and model parameters $\Theta = [\{\mu_l, \kappa_l\}_{l=1}^{L}]$ is given by $P(X|Y, \Theta) \propto \prod_{i=1}^{M} p(y_i|x_i)P(X)$
- The posterior can hence be written as $P(X|Y, \Theta) = \frac{1}{Z'} e^{-(U_s(X) + U_l(X) + U_D(Y))}$, where $U_D(Y) = -\sum_{i=1}^{M} log(p(y_i|x_i))$

Notation for this problem

- $U_s(X) = \sum_{i=1}^{M} \sum_{j \in N_i} V(X_i, X_j)$ where the potts potential function is defined as $V(X_i, X_j) = \beta_s(1 - \delta(X_i - X_j))$, where $\delta(z) = 1$ if $z = 0$ and $\delta(z) = 0$ otherwise
- The label cost $U_l(X) = \beta_l \sum_{i=1}^{L} \delta_p(X)$, where $\delta_p(X) = 1$, if $\exists$ i such that $X_i = p$ and 0 otherwise
- The posterior distribution of X given the data Y and model parameters $\Theta = [\{\mu_l, \kappa_l\}_{l=1}^{L}]$ is given by $P(X|Y, \Theta) \propto \prod_{i=1}^{M} p(y_i|x_i)P(X)$
- The posterior can hence be written as $P(X|Y, \Theta) = \frac{1}{Z'} e^{-(U_s(X)+U_l(X)+U_D(Y))}$, where $U_D(Y) = -\sum_{i=1}^{M} log(p(y_i|x_i))$
- Henceforth, we intend to maximise this posterior probability over the labels and parameters. This is same as minimising the total energy function $U = U_s(X) + U_l(X) + U_D(Y)$. Smell of a graph cut?

Notation for this problem

- $U_s(X) = \sum_{i=1}^{M} \sum_{j \in N_i} V(X_i, X_j)$ where the potts potential function is defined as $V(X_i, X_j) = \beta_s(1 - \delta(X_i - X_j))$, where $\delta(z) = 1$ if $z = 0$ and $\delta(z) = 0$ otherwise
- The label cost $U_l(X) = \beta_l \sum_{i=1}^{L} \delta_p(X)$, where $\delta_p(X) = 1$, if $\exists$ i such that $X_i = p$ and 0 otherwise
- The posterior distribution of X given the data Y and model parameters $\Theta = [\{\mu_l, \kappa_l\}_{l=1}^{L}]$ is given by $P(X|Y, \Theta) \propto \prod_{i=1}^{M} p(y_i|x_i)P(X)$
- The posterior can hence be written as $P(X|Y, \Theta) = \frac{1}{Z'} e^{-(U_s(X)+U_l(X)+U_D(Y))}$, where $U_D(Y) = -\sum_{i=1}^{M} log(p(y_i|x_i))$
- Henceforth, we intend to maximise this posterior probability over the labels and parameters. This is same as minimising the total energy function $U = U_s(X) + U_l(X) + U_D(Y)$. Smell of a graph cut?
- $\hat{X} = arg\ min_X U$

Parameters and Labels Update

VMF Parameter Estimate

To get the optimal parameters, we maximise the likelihood. We get $\overline{r_l} = \frac{\sum_{\{i|X(i)=l\}} y_i}{N}$, $\hat{\mu}_l = \frac{\overline{r_l}}{\|\overline{r_l}\|}$ and $\hat{\kappa}_l = \frac{\overline{r_l}T - \overline{r_l}^3}{1-\overline{r_l}^2}$

Parameters and Labels Update

VMF Parameter Estimate

To get the optimal parameters, we maximise the likelihood. We get
$\overline{r_l} = \frac{\sum_{\{i|X(i)=l\}} y_i}{N}$, $\hat{\mu}_l = \frac{\overline{r_l}}{\|\overline{r_l}\|}$ and $\hat{\kappa}_l = \frac{\overline{r_l}T - \overline{r_l}^3}{1-\overline{r_l}^2}$

α Expansion Algorithm

1. Start with an arbitrary labelling $\mathbf{f}$
2. Set **success** $:= \mathbf{0}$
3. For each label $\alpha \in \mathcal{L}$
    3.1 Find $\hat{\mathbf{f}} = \arg\min \mathbf{E}(\mathbf{f}')$ among $\mathbf{f}'$ within one α expansion of α
    3.2 If $\mathbf{E}(\hat{\mathbf{f}}) < \mathbf{E}(\mathbf{f})$, set $\mathbf{f} = \hat{\mathbf{f}}$ and success:=1
4. If **success** $= \mathbf{1}$ goto 2
5. Return $\mathbf{f}$

Table of Contents

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters
- Second we generated voxel time series in each cluster such that voxels correlations within a cluster ρ_{in} will be different from the voxel correlation across the clusters ρ_{ac}. We generated datasets with different values of ρ_{in} and ρ_{ac}

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters
- Second we generated voxel time series in each cluster such that voxels correlations within a cluster ρ_{in} will be different from the voxel correlation across the clusters ρ_{ac}. We generated datasets with different values of ρ_{in} and ρ_{ac}
    1. High correlation within the cluster ($\rho_{in} = 0.7$) and low correlation across the clusters ($\rho_{ac} = 0.3$)

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters
- Second we generated voxel time series in each cluster such that voxels correlations within a cluster ρ_{in} will be different from the voxel correlation across the clusters ρ_{ac}. We generated datasets with different values of ρ_{in} and ρ_{ac}
    1. High correlation within the cluster ($\rho_{in} = 0.7$) and low correlation across the clusters ($\rho_{ac} = 0.3$)
    2. High correlation within the cluster ($\rho_{in} = 0.7$) and across the clusters ($\rho_{ac} = 0.6$)

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters
- Second we generated voxel time series in each cluster such that voxels correlations within a cluster ρ_{in} will be different from the voxel correlation across the clusters ρ_{ac}. We generated datasets with different values of ρ_{in} and ρ_{ac}
  1. High correlation within the cluster ($\rho_{in} = 0.7$) and low correlation across the clusters ($\rho_{ac} = 0.3$)
  2. High correlation within the cluster ($\rho_{in} = 0.7$) and across the clusters ($\rho_{ac} = 0.6$)
  3. Medium correlation within the cluster ($\rho_{in} = 0.5$) and low correlation across the clusters ($\rho_{ac} = 0.2$)

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters
- Second we generated voxel time series in each cluster such that voxels correlations within a cluster ρ_{in} will be different from the voxel correlation across the clusters ρ_{ac}. We generated datasets with different values of ρ_{in} and ρ_{ac}
    1. High correlation within the cluster ($\rho_{in} = 0.7$) and low correlation across the clusters ($\rho_{ac} = 0.3$)
    2. High correlation within the cluster ($\rho_{in} = 0.7$) and across the clusters ($\rho_{ac} = 0.6$)
    3. Medium correlation within the cluster ($\rho_{in} = 0.5$) and low correlation across the clusters ($\rho_{ac} = 0.2$)
    4. Medium correlation within the cluster ($\rho_{in} = 0.5$) and across the clusters ($\rho_{ac} = 0.4$)

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters
- Second we generated voxel time series in each cluster such that voxels correlations within a cluster ρ_{in} will be different from the voxel correlation across the clusters ρ_{ac}. We generated datasets with different values of ρ_{in} and ρ_{ac}
  1. High correlation within the cluster ($\rho_{in} = 0.7$) and low correlation across the clusters ($\rho_{ac} = 0.3$)
  2. High correlation within the cluster ($\rho_{in} = 0.7$) and across the clusters ($\rho_{ac} = 0.6$)
  3. Medium correlation within the cluster ($\rho_{in} = 0.5$) and low correlation across the clusters ($\rho_{ac} = 0.2$)
  4. Medium correlation within the cluster ($\rho_{in} = 0.5$) and across the clusters ($\rho_{ac} = 0.4$)
  5. Low correlation within the cluster ($\rho_{in} = 0.3$) and very low correlation across the clusters ($\rho_{ac} = 0.1$)

Simulated data using computer based simulations

- First we synthesize the cluster shapes and sizes using Gibbs sampling procedure wherein we impose spatial regularization using MRF to simulate spatially smooth clusters
- Second we generated voxel time series in each cluster such that voxels correlations within a cluster ρ_{in} will be different from the voxel correlation across the clusters ρ_{ac}. We generated datasets with different values of ρ_{in} and ρ_{ac}
  1. High correlation within the cluster ($\rho_{in} = 0.7$) and low correlation across the clusters ($\rho_{ac} = 0.3$)
  2. High correlation within the cluster ($\rho_{in} = 0.7$) and across the clusters ($\rho_{ac} = 0.6$)
  3. Medium correlation within the cluster ($\rho_{in} = 0.5$) and low correlation across the clusters ($\rho_{ac} = 0.2$)
  4. Medium correlation within the cluster ($\rho_{in} = 0.5$) and across the clusters ($\rho_{ac} = 0.4$)
  5. Low correlation within the cluster ($\rho_{in} = 0.3$) and very low correlation across the clusters ($\rho_{ac} = 0.1$)
  6. Low correlation within cluster ($\rho_{in} = 0.3$) and across the clusters ($\rho_{ac} = 0.2$)

Results on Simulated Data

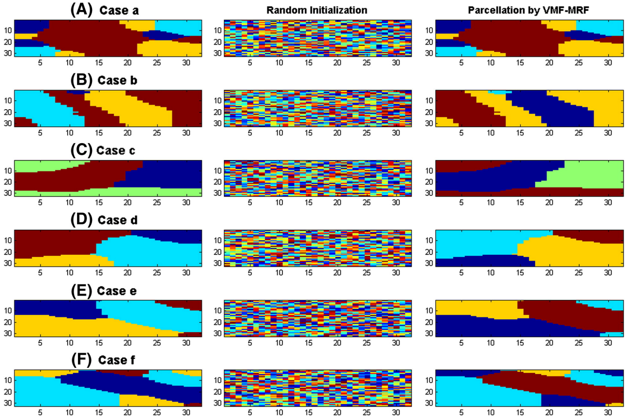


Figure: All 6 Simulated data samples

Iterations in case 4

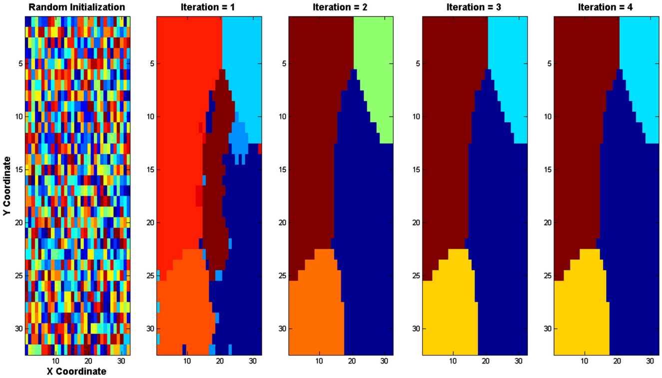


Figure: Clusters estimated by this algorithm in case 4

Clustering Human Brain RS-fMRI Data

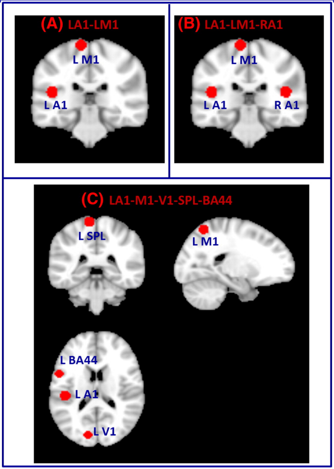


Figure: Clusters constructed from fMRI data(3 Input Data Sets)

Clustering Human Brain RS-fMRI Data

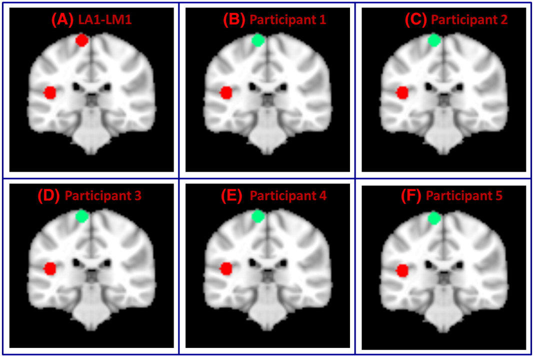


Figure: Left auditory and motor cortices (LA1-LM1)

Clustering Human Brain RS-fMRI Data

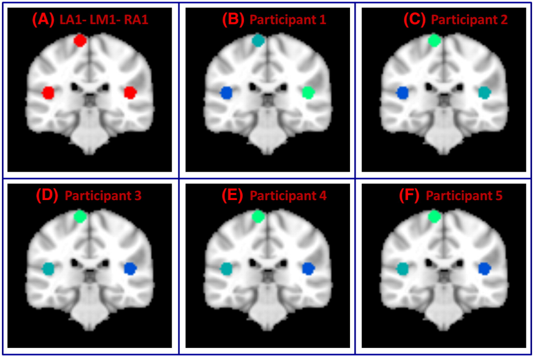


Figure: Left auditory, motor and right auditory cortices(LA1-LM1-RA1)

Clustering Human Brain RS-fMRI Data

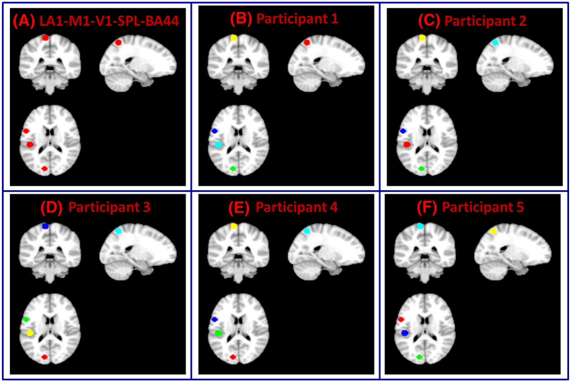


Figure: Left auditory, motor, and visual cortices, the superior parietal lobule (SPL) and inferior frontal gyrus (Brodmann Area BA 44) (LA1-M1-V1-SPL-BA44)

Inferences

- Cases 1, 3 and 5 involve high spatial contrast ($\rho_{in} > \rho_{ac}$) whereas cases 2, 4 and 6 involve low spatial contrast ($\rho_{in} \approx \rho_{ac}$) and should be more difficult to parcellate. We evaluated the robustness of our method using these test cases and as visible the algorithm correctly parcellated the data sets.

Inferences

- Cases 1, 3 and 5 involve high spatial contrast ($\rho_{in} > \rho_{ac}$) whereas cases 2, 4 and 6 involve low spatial contrast ($\rho_{in} \approx \rho_{ac}$) and should be more difficult to parcellate. We evaluated the robustness of our method using these test cases and as visible the algorithm correctly parcellated the data sets.
- It predicted the number of cluster labels and also parcellated spatially contiguous as well as non-contiguous regions appropriately(even in case 6)

Inferences

- Cases 1, 3 and 5 involve high spatial contrast ($\rho_{in} > \rho_{ac}$) whereas cases 2, 4 and 6 involve low spatial contrast ($\rho_{in} \approx \rho_{ac}$) and should be more difficult to parcellate. We evaluated the robustness of our method using these test cases and as visible the algorithm correctly parcellated the data sets.
- It predicted the number of cluster labels and also parcellated spatially contiguous as well as non-contiguous regions appropriately(even in case 6)
- The usage of α expansion significantly improved the running time over ICM or simulated annealing based approaches. It does so by allowing updates on several voxels simultaneously instead of just one as in other cases.

Inferences

- Cases 1, 3 and 5 involve high spatial contrast ($\rho_{in} > \rho_{ac}$) whereas cases 2, 4 and 6 involve low spatial contrast ($\rho_{in} \approx \rho_{ac}$) and should be more difficult to parcellate. We evaluated the robustness of our method using these test cases and as visible the algorithm correctly parcellated the data sets.
- It predicted the number of cluster labels and also parcellated spatially contiguous as well as non-contiguous regions appropriately(even in case 6)
- The usage of α expansion significantly improved the running time over ICM or simulated annealing based approaches. It does so by allowing updates on several voxels simultaneously instead of just one as in other cases.
- Now, since this approach worked for cases where the ground truth was already known(as depicted in the above results), we can run this algorithm to develop finer insights into functional brain segmentation

What can we expect more?

- To observe the sensitivity of the method to the factors β_s and β_I,we could examine the outputs generated for different combinations of β_s and β_I.

What can we expect more?

- To observe the sensitivity of the method to the factors β_s and β_l,we could examine the outputs generated for different combinations of β_s and β_l.
- For a given β_s, β_l pair, we can also study the dependence on the initial random clustering by creating a connectivity matrix for each initialization and using them to build a consensus matrix.

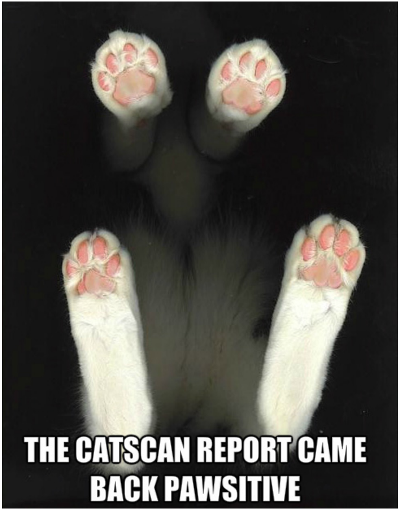
What can we expect more?

- To observe the sensitivity of the method to the factors β_s and β_l,we could examine the outputs generated for different combinations of β_s and β_l.
- For a given β_s, β_l pair, we can also study the dependence on the initial random clustering by creating a connectivity matrix for each initialization and using them to build a consensus matrix.
- We expect the consensus matrix to average out the opinions of all the initializations for all voxel pairs and hence be a bimodal distribution with modes at 0 and 1.

What can we expect more?

- To observe the sensitivity of the method to the factors β_s and β_I,we could examine the outputs generated for different combinations of β_s and β_I.
- For a given β_s, β_I pair, we can also study the dependence on the initial random clustering by creating a connectivity matrix for each initialization and using them to build a consensus matrix.
- We expect the consensus matrix to average out the opinions of all the initializations for all voxel pairs and hence be a bimodal distribution with modes at 0 and 1.
- We could also generate updates for β_s and β_I, which may be done by several methods such as EM optimization with priors on β_s and β_I having latent variables.

In a Lighter Vein



Bibliography

- For the research paper click here
- For α expansion click here
- For Von Mises-Fisher Distribution click here

Thanks for your attention! Any questions?

Hope you slept comfortably!