Mixture of multivariate multiple-scaled Student distributions : application to the characterization of brain tumors by multiparametric MRI.

Alexis Arnaud

Supervisors : F. Forbes, E. Barbier, B. Lemasson
Laboratories : LJK, INRIA & GIN
Funding : UJF

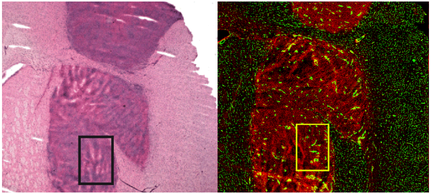
1st year-PhD
26 June 2015

Outline

Outline

How to characterize brain tumor ?

Histology vs multiparametric MRI

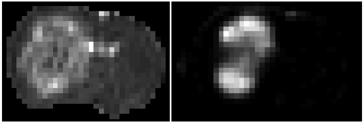


Histology section

- + gold standard : provides the ground truth
- + precise information
- – local information : only parts of the tumor may be sampled
- – invasive operation (biopsy), not always feasible

How to characterize brain tumor ?

Histology vs multiparametric MRI



2 MRI maps

- − under development method
- + global information : whole tumor visible
- + non-invasive operation

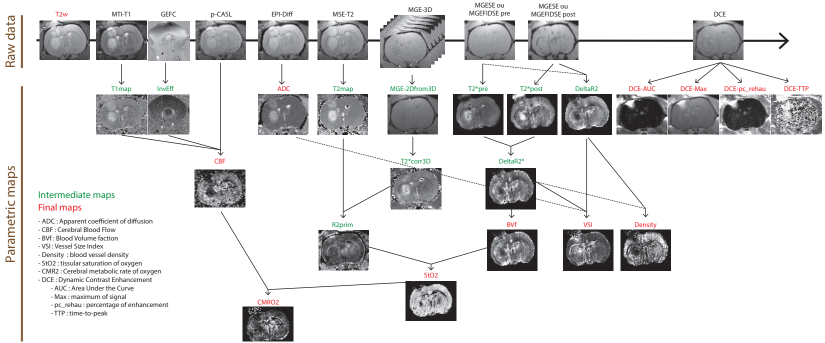
Goal : find the voxels inside the MRI maps which belong to the tumor, in order to characterize a tumor, to avoid invasive biopsies.

Alexis Arnaud
Supervisors : F. Forbes, E. Barbier, B. Lemasson *Laboratories* : LJK, INRIA & GI

How to characterize brain tumor ?

Problem of multiparametric MRI

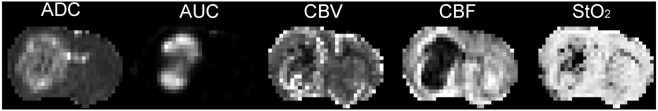
How to extract information from all of the parametric maps ?



- **Approach** : multivariate clustering with mixture models.

Outline

Multiparametric MRI data



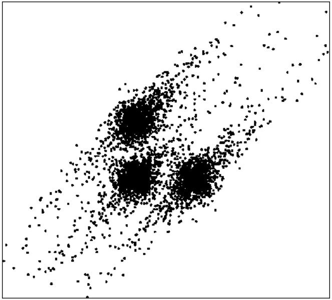
- 5 physiological parameters :
  - **ADC** : apparent diffusion coefficient
  - **CBV** : cerebral blood volume
  - **CBF** : cerebral blood flow
  - **AUC** : blood vessel permeability
  - **StO_2** : oxygen saturation

- 5 dimensional data set :

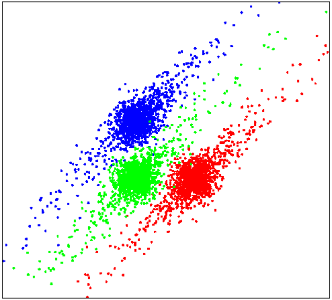
$\boldsymbol{Y} = \{\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{\mathrm{N}}\}$ the set of all voxels, of size N
with $\boldsymbol{Y}_n = \{\boldsymbol{Y}_{n,\mathrm{ADC}}, \ldots, \boldsymbol{Y}_{n,\mathrm{StO_2}}\}$ the mesures on the nth voxel

Clustering of MRI data

We supose that data arise from K different classes (for K different tissues), and we want to recover those classes.



Simulated data



Latent classification

Clustering of MRI data via mixture modeling

Let $\boldsymbol{Z}$ be the latent variable which links one observation to one class :

$$\left\{ \begin{array}{ccc} (\boldsymbol{Y}_n \,|\, \boldsymbol{Z}_n = k) & \sim & f_k\left(\boldsymbol{\theta}_k\right) \\ \boldsymbol{Z}_n & \sim & \mathcal{M}\left(\pi_1, \ldots, \pi_{\mathrm{K}}\right) \end{array} \right. \quad (1)$$

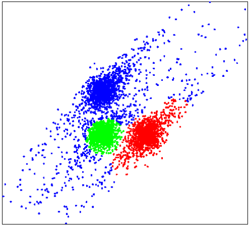
$\Leftrightarrow$

$$\Pr\left(\boldsymbol{y}_n \,;\, \boldsymbol{\theta}\right) = \sum_{k=1}^{\mathrm{K}} \pi_k f_k\left(\boldsymbol{y}_n \,;\, \boldsymbol{\theta}_k\right) \quad (2)$$

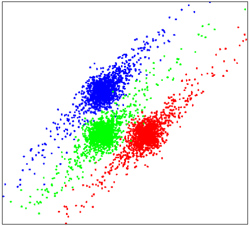
Clustering of MRI data via mixture modeling

In a previous study, $f_k(\boldsymbol{\theta}_k) = \mathcal{N}_5(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ with $\boldsymbol{\mu}_k \in \mathbb{R}^5$ and $\boldsymbol{\Sigma}_k \in \mathcal{S}^+_{5\times 5}(\mathbb{R})$: lack of flexibility in cluster shape modeling (Coquery *et al.* - 2014).

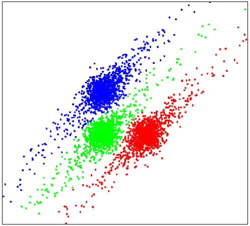
In this study, $f_k(\boldsymbol{\theta}_k)$ is a heavy-tailed distribution : a multivariate multiple-scaled Student distribution (MMSD).



Gaussian



Simulation



MMSD

Outline

Standard multivariate Student distribution

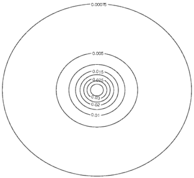
One possible form of a M-dimensional distribution

$$p_{\mathrm{MS}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu},\boldsymbol{\Sigma},\nu\right) \quad = \quad \frac{\Gamma\left(\frac{\nu+\mathrm{M}}{2}\right)}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}\,\Gamma\left(\frac{\nu}{2}\right)(\pi\nu)^{\frac{\mathrm{M}}{2}}}\left[1+\frac{\delta^2\left(\boldsymbol{y},\boldsymbol{\mu},\boldsymbol{\Sigma}\right)}{\nu}\right]^{-\frac{\nu+\mathrm{M}}{2}}$$

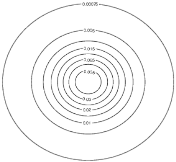
with : $\delta^2\left(\boldsymbol{y},\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = (\boldsymbol{y}-\boldsymbol{\mu})^{\mathrm{t}}\,\boldsymbol{\Sigma}^{-1}\,(\boldsymbol{y}-\boldsymbol{\mu})$

the Mahalanobis distance

- But the degree of freedom remains scalar.



$\nu = 5$



$\nu = 10$

Standard multivariate Student distribution

Useful representation : infinite mixture of scaled Gaussians

$$p_{\mathrm{MS}}\left(\boldsymbol{y} \; ; \; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu\right) = \frac{\Gamma\left(\frac{\nu+\mathrm{M}}{2}\right)}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}\,\Gamma\left(\frac{\nu}{2}\right)(\pi\nu)^{\frac{\mathrm{M}}{2}}}\left[1+\frac{\delta^2\left(\boldsymbol{y},\boldsymbol{\mu},\boldsymbol{\Sigma}\right)}{\nu}\right]^{-\frac{\nu+\mathrm{M}}{2}}$$

$$= \int_0^{\infty}\mathcal{N}_{\mathrm{M}}\left(\boldsymbol{y} \; ; \; \boldsymbol{\mu}, \frac{1}{w}\boldsymbol{\Sigma}\right)\mathcal{G}\left(w \; ; \; \frac{\nu}{2}, \frac{\nu}{2}\right)\mathrm{d}w$$

with :

- $\mathcal{N}_{\mathrm{M}}$ the M-multivariate Gaussian distribution
- $\mathcal{G}$ the Gamma distribution
- the real latent variable W is called the weight

Multivariate multiple-scaled Student distribution

Use of the eigenvalue decomposition to get a multidimensional degree of freedom (Forbes and Wraith - 2014)

Let U and D the eigenvalue decomposition of covariance matrix :
$\boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{\mathrm{t}}$, with :

- $\boldsymbol{U} \in \mathcal{O}(\mathrm{M})$ the orthogonal matrix of eigenvectors,
- $\boldsymbol{D} \in \mathcal{D}(\mathrm{M})$ the diagonal matrix of eigenvalues.

$$p_{\mathrm{MS}}(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \boldsymbol{U}, \boldsymbol{D}, \nu) = \int_{\mathbb{R}_+^*} \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{y}\ ;\ \boldsymbol{\mu}, \frac{1}{w}\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{\mathrm{t}}\right) \mathcal{G}\left(w\ ;\ \frac{\nu}{2}, \frac{\nu}{2}\right)\, \mathrm{d}w$$

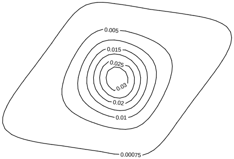
Multivariate multiple-scaled Student distribution

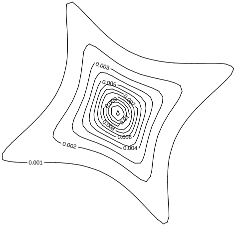
Use of the eigenvalue decomposition to get a multidimensional degree of freedom (Forbes and Wraith - 2014)

Let $\boldsymbol{W} \in \mathbb{R}_{+*}^{\mathrm{M}}$ a M-dimensional weight (one scalar weight for each dimension) :

$$p_{\mathrm{MMS}}\left(\boldsymbol{y} \; ; \boldsymbol{\mu}, \boldsymbol{U}, \boldsymbol{D}, \boldsymbol{\nu}\right) =$$

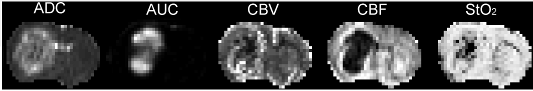
$$\int_{\mathbb{R}_{+*}^{\mathrm{M}}} \mathcal{N}_{\mathrm{M}}\left(\boldsymbol{y} \; ; \boldsymbol{\mu}, \boldsymbol{U} \mathrm{diag}\left(\mathbf{w}\right)^{-1} \boldsymbol{D} \boldsymbol{U}^{\mathrm{t}}\right) \prod_{m=1}^{\mathrm{M}} \mathcal{G}\left(w_m \; ; \frac{\nu_m}{2}, \frac{\nu_m}{2}\right) \, \mathrm{d}\mathbf{w}$$





Outline

Multiparametric MRI data

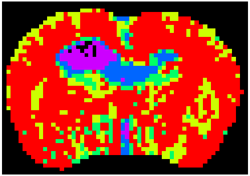


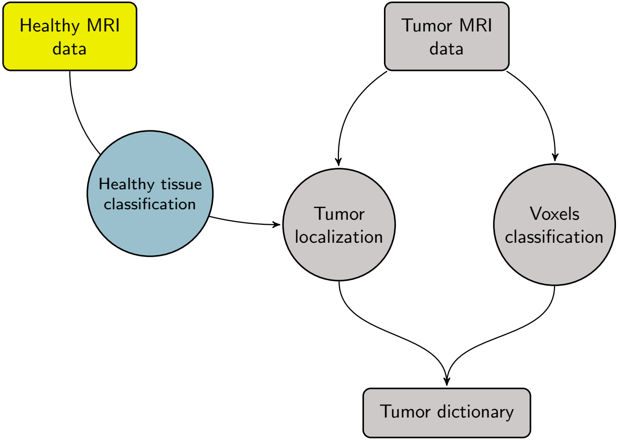
- **26 parameters per class**
- estimation by **EM** and **Flury & Gautschi** or **Stiefel manifold optimization** algorithms :
  - writen in C++
  - use inside R with the **Rcpp package**
- choice of the number of classes according with **BIC** and **ICL** criterions :
  - 10 to 50 repetitions in parallel
  - use of the **snow package**

- data in **dimension 5**
- 8 healthy **rats** (49 000 voxels)
- 37 **rats** with tumors (290 000 voxels)
- 4 tumor models : **9L**, **C6**, **F98**, **RG2**

Processing pipeline

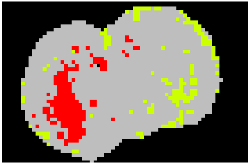
Healthy voxels classification

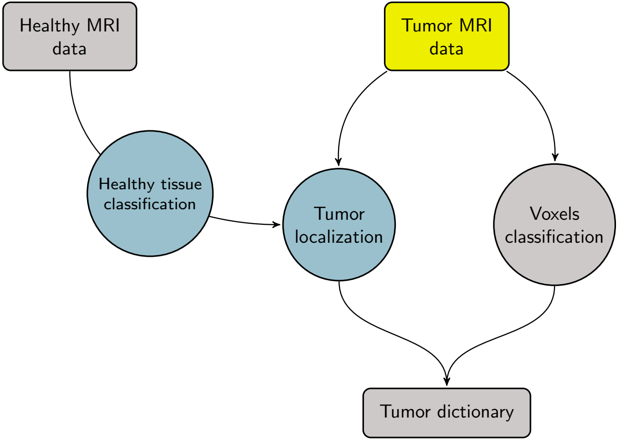




Processing pipeline

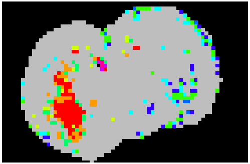
Tumor localization (ROI)

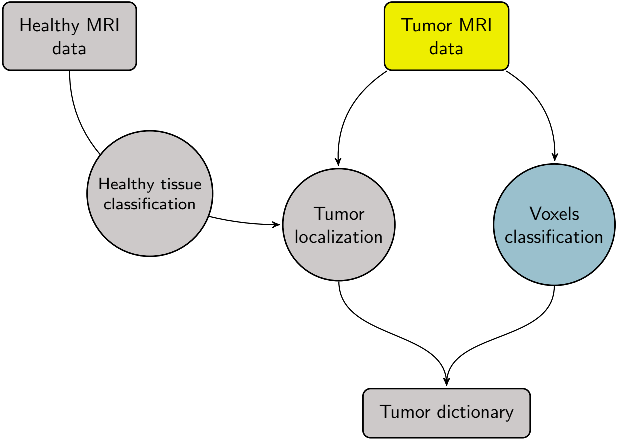




Processing pipeline

Voxels classification

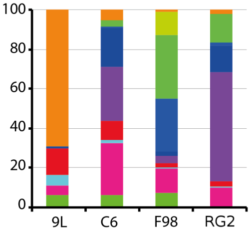


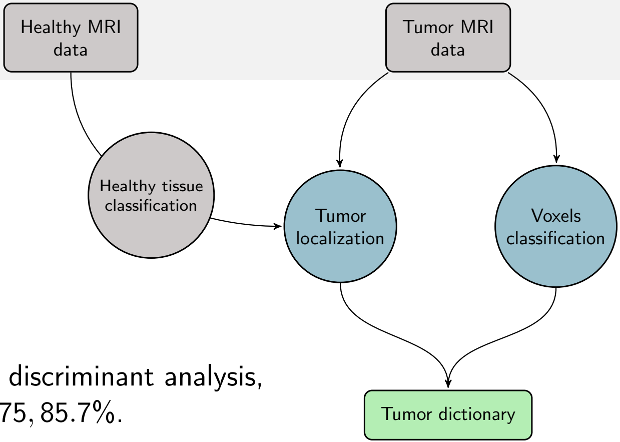


Alexis Arnaud
Supervisors : F. Forbes, E. Barbier, B. Lemasson Laboratories : LJK, INRIA & GI

Processing pipeline

Tumor dictionary





Classification rate of discriminant analysis, by tumor : $50, 62.5, 75, 85.7\%$.

▶ **Except for the 9L rats, the classification rates are as good as a previous study with a Gaussian model and a manual tumor delimitation.**

Outline

Work in progress

- Validation of the protocol.
- Taking into account spatial dependences (Markov field).
- Automatic selection of the number of classes (Bayesian extension).
- Parameters sensitivity analysis.
- Link between histology and automatic tissue characterization.

Bibliography

- **Coquery N, Francois O, Lemasson B, Debacker C, Farion R, Rémy C, et Barbier E (2014)**, *Microvascular MRI and unsupervised clustering yields histology-resembling images in two rat models of glioma*, Journal of Cerebral Blood Flow & Metabolism, volume 34, number 8, 1354–62.

- **Forbes F, et Wraith D (2014)**, *A new family of multivariate heavy-tailed distributions with variable marginal amounts of tailweights: Application to robust clustering*, Statistics and Computing, volume 24, number 6, 971–984.

The end

Thank you !

Image credit : dianephotos.deviantart.com