

micro et nanoélectronique  
microsystèmes  
intelligence ambiante  
biologie et santé  
chaîne de l'image



## Riemannian geometry applied to BCI classification

Barachant Alexandre et al.

October 14, 2010

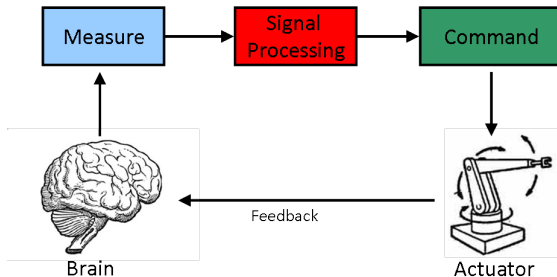
leti

MINATEC<sup>®</sup>

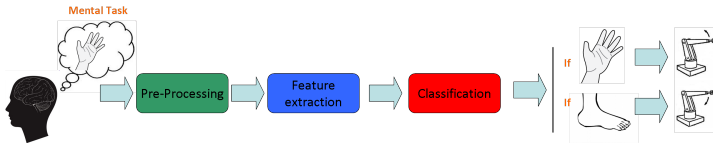
INSTITUT  
CARNOT  
CEA-LETI



# The Brain Computer Interface cycle

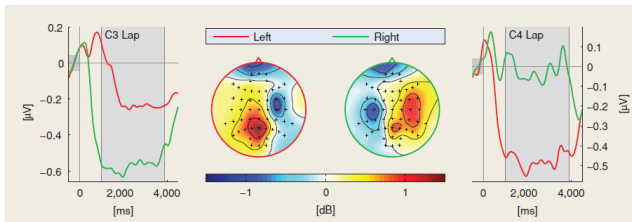


- Goal of signal processing : detect some specific brain activity within the whole brain activity.



# Covariance Matrices in BCI

- The most common way to generate specific brain activity is motor imagery. For example, imagination of right hand movement.
- Motor imagery induces a spatially localised change in a specific frequency band.

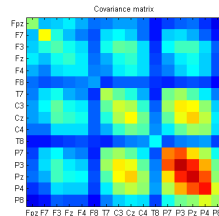
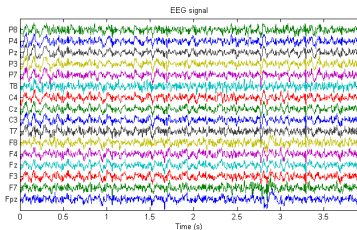


**Figure:** 9-13Hz band power change for a right and left hand movement [2]

# Covariance Matrices in EEG BCI

- The spatial covariance matrices of EEG signal are a natural choice since they contain spatial and power informations.
- We denote by  $\mathbf{E} \in R^{n \times t}$  a given mean centred EEG recording epoch with  $n$  electrodes and  $t$  samples. The spatial sample covariance matrix  $\mathbf{P}$  is given by :

$$\mathbf{P} = \frac{1}{t-1} \mathbf{E} \mathbf{E}^T$$

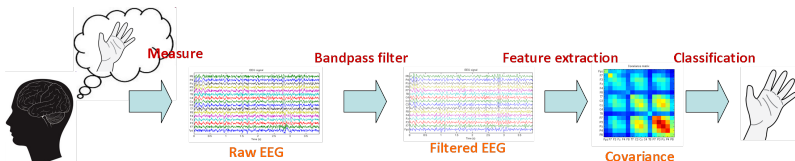


# Riemannian space

- Covariance matrices are Symmetric and Positive Definite (SPD) matrices.
- SPD matrices lie in a Riemannian space with the following metric (local at  $P$ ):

$$ds^2 = \|P^{-1/2} dP P^{-1/2}\|_F^2$$

- Using a differential geometry framework, we propose algorithms in order to classify covariance matrices in their native space.



© CEA 2008. Tous droits réservés.  
Toute reproduction totale ou partielle sans que les supports ou les utilisations de contenu de ce document est interdite sans l'autorisation écrite préalable de CEA.  
All rights reserved. Any reproduction in whole or in part without the express written consent of CEA is prohibited.

# Differential geometry framework for SPD matrices

- Geodesic between two SPD matrices ( $P_1$  and  $P_2$ ):

$$\gamma(t) = P_1^{1/2} \left( P_1^{-1/2} P_2 P_1^{-1/2} \right)^t P_1^{1/2}$$

with  $t \in [0 : 1]$

- Distance between two SPD matrices ( $P_1$  and  $P_2$ ):

$$\delta_R(P_1, P_2) = \int_0^1 \gamma(t) dt = \|\text{Log}(P_1^{-1} P_2)\|_F$$

- Mean of  $K$  SPD matrices :

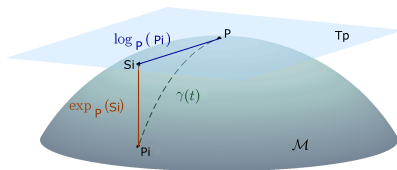
$$\mathfrak{G}(P_1, \dots, P_K) = \underset{P \in P(n)}{\text{argmin}} \sum_{k=1}^K \delta_R^2(P, P_k)$$

No closed form expression, need an iterative algorithm [3].

© CEA 2009. Tous droits réservés.  
Toute reproduction totale ou partielle sans que les supports ou les utilisations de contenu de ce document est interdite sans l'autorisation écrite préalable de CEA.  
All rights reserved. Any reproduction in whole or in part without the express written consent of CEA is prohibited.

# Tangent space

- Given a point  $\mathbf{P} \in P(n)$ , it is possible for every point  $\mathbf{P}_i \in P(n)$ , to identify a tangent vector  $\mathbf{S}_i \in S(n)$  such as  $\mathbf{S}_i = \dot{\gamma}(0)$  with  $\gamma(t)$  the geodesic between  $\mathbf{P}$  and  $\mathbf{P}_i$ .
- The whole set of tangent vectors define the Tangent Space, which is **Euclidean**.
- Log map :  $\mathbf{S}_i = \text{Log}_{\mathbf{P}}(\mathbf{P}_i) = \mathbf{P}^{1/2} \text{Log}(\mathbf{P}^{-1/2} \mathbf{P}_i \mathbf{P}^{-1/2}) \mathbf{P}^{1/2}$
- Exp map :  $\mathbf{P}_i = \text{Exp}_{\mathbf{P}}(\mathbf{S}_i) = \mathbf{P}^{1/2} \text{Exp}(\mathbf{P}^{-1/2} \mathbf{S}_i \mathbf{P}^{-1/2}) \mathbf{P}^{1/2}$



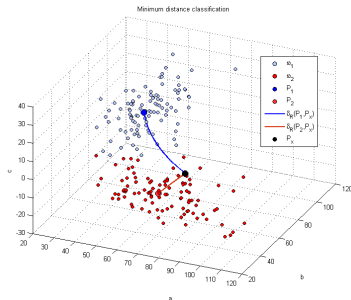
**Figure:** Tangent space of the manifold  $\mathcal{M}$  at point  $\mathbf{P}$ ,  $\mathbf{S}_i$  the tangent vector of  $\mathbf{P}_i$  and  $\gamma(t)$  the geodesic between  $\mathbf{P}$  and  $\mathbf{P}_i$ .

# Minimum Distance to Mean (MDM) classifier

- Use Riemann Distance and Geometric mean to classify covariance matrices.

- $P_1 = \mathcal{G}(P_i \rightarrow \omega_i = 1)$
- $P_2 = \mathcal{G}(P_i \rightarrow \omega_i = 2)$
- For a given matrix  $P_x$  of unknown class  $\omega_x$  :

$$\omega_x = \underset{i}{\operatorname{argmin}}(\delta_R(P_x, P_i))$$





# Geodesic filtering and classification in Tangent Space

- When previous algorithms are not enough powerful, tangent space can be used in order to manipulate data (filtering or classification).
- Data are mapped to the tangent space using Logarithmic map

$$\forall i, \mathbf{S}_i = \text{Log}_{\mathbf{P}}(\mathbf{P}_i)$$

with  $\mathbf{P}$  the Geometric mean of all data.

- Since Tangent space are Euclidean, we can treat the  $\mathbf{S}_i$  as vectors (taking care of removing symmetric elements)

$$\tilde{\mathbf{S}}_i = \text{vect}(\mathbf{S}_i)$$

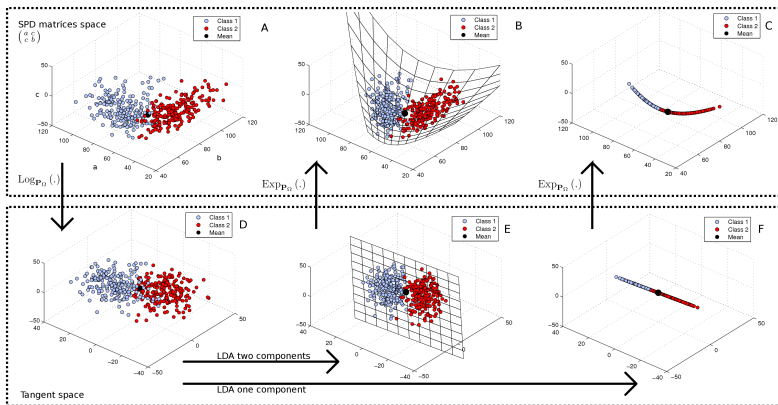
- We can apply classical methods on the  $\tilde{\mathbf{S}}_i$  and if necessary, go back to the Riemannian Space.

$$\tilde{\mathbf{P}}_i = \text{Exp}_{\mathbf{P}}(\text{unvect}(\tilde{\mathbf{S}}_i))$$

© CEA 2008. Tous droits réservés.  
Toute reproduction totale ou partielle sans que les supports de ce document est interdite sans l'autorisation écrite préalable de CEA.  
All rights reserved. Any reproduction in whole or in part without the written consent of CEA is prohibited.

# Example : Geodesic Filtering with Fisher discriminant analysis (FGDA)

- Using Fisher LDA in tangent space, we apply a geodesic filtering in order to enhance the discrimination of data.



# Results

- Comparison of our methods with an implementation of the reference method [2]. (fréquential filters, spatial filters and LDA classification)
- Two datasets
  - Dataset without artefacts : Dataset IVa of BCI competition III <sup>1</sup>. Five subject, 118 electrodes (we use a subset of 9), two motor imagery classes (right hand and right foot) and 280 trials.
  - Dataset with artefacts : Our own dataset. 9 users, 16 electrodes, three classes (hand, foot and rest) and 180 trials.

<sup>1</sup><http://www.bbc.de/competition/iii/>

# Results

- Competition BCI dataset, classification error rate using Cross-validation.

User	Reference	MDM	FGDA
aa	26	28.9	22.5
al	3.2	3.9	2.8
av	34.2	39.6	34.2
aw	6.4	11	7.4
ay	7.4	12.1	7.1
Mean	$15.5 \pm 13.8$	$19.1 \pm 14.7$	$14.8 \pm 13.6$

- Results are similar to results obtained with the reference method.
- MDM algorithm is very simple but offers good results.
- Large inter-subjects variability, a common issue in BCI.

- Our dataset, correct classification rate.

User	Reference	MDM
UserG	34, 9	55, 0
UserA	48, 3	61, 6
UserH	44, 7	44, 1
UserI	73, 4	76, 0
UserB	50, 0	58, 4
UserJ	38, 5	67, 1
UserK	31, 5	37, 0
UserL	55, 5	58, 7
UserE	65, 4	57, 7
mean	49, 1	57, 3

- Results are significantly better than the reference method (Riemannian methods are robust)
- FGDA can not be applied, there is not enough data. 180 trials for a Tangent space of dimension 136 (16 electrodes).
- Need regularised FGDA.

# Conclusion

- Covariances matrices can be directly classified using a Riemannian framework.
- Distance-based classification methods (MDM, kNN, k-means, ...) can be directly transposed in Riemannian Space.
- More sophisticated methods (LDA, SVM, NN, ...) can be transposed using Tangent Space.
- Tangent space is high dimensional ( $\frac{n \times (n+1)}{2}$ ), regularisation is often needed.
- No more need of spatial filtering (classical way to use spatial information in BCI)
- Riemannian methods are robust.

# Perspectives

- Regularisation in Tangent space.
- Dimensional reduction.
- Use these concepts in other kind of BCI.

# References



**BCI competition III, dataset IVa.**

[http://ida.first.fhg.de/projects/bci/competition\\_iii](http://ida.first.fhg.de/projects/bci/competition_iii).



**B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller.**

Optimizing spatial filters for robust EEG Single-Trial analysis.

*Signal Processing Magazine, IEEE*, 25(1):41–56, 2008.



**Maher Moakher.**

A differential geometric approach to the geometric mean of symmetric Positive-Definite matrices.

*SIAM J. Matrix Anal. Appl.*, 26(3):735–747, 2005.