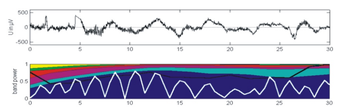
Clustering of EEG-Segments using Hierarchical Agglomerative Methods and Self-Organized Maps

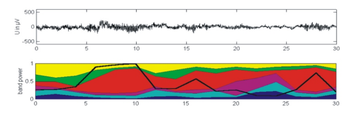
David Sommer, Martin Golz

Abstract. EEG segments recorded during microsleep events were transformed to the frequency domain and were subsequently clustered without the common summation of the power densities in spectral bands. Any knowledge about the number of clusters doesn't exist. The hierarchical agglomerative clustering procedures were terminated with several standard measures of intracluster and intercluster variances. The results were inconsistently. The winner histogram of Self-organizing maps showed also no evidence. The analysis of the U-matrix together with the watershed transform, a method from image processing, resulted in separable clusters. Like in many other procedures the number of clusters is determined with one threshold parameter. The proposed method is working fully automatically.

Introduction. Slow eye movements (SEM):typical patterns of eye movements during the transition from wakefulness to sleep
Sleepy drivers in a truck driving simulation task sometimes also show SEM immediately before or during a microsleep episode
EEG - segments (2 sec) --> spectral power densities [2 .. (0.5) .. 25 Hz] --> cluster analysis
Question: what is the typical spectral characteristic of an EEG record during a SEM?



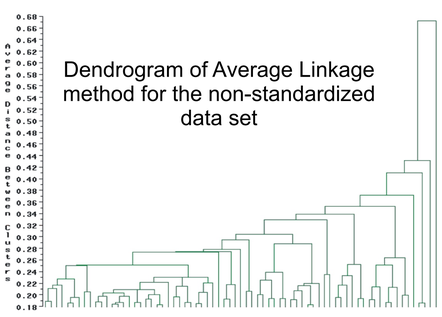
upper chart: EOG of a typical SEM event
lower chart: relative spectral power density



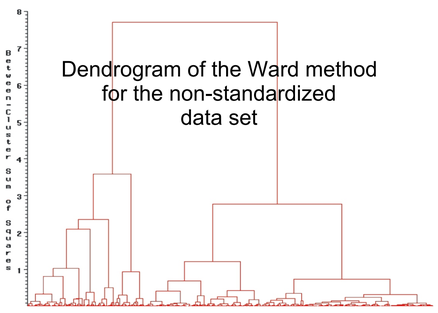
upper chart: EEG of a typical SEM event
lower chart: relative spectral power density

Clustering with Agglomerative Hierarchical Methods

Estimated numbers of clusters obtained from five different agglomerative hierarchical methods with six different measures



Dendrogram of Average Linkage method for the non-standardized data set



Dendrogram of the Ward method for the non-standardized data set

Measure	Method				
	Single Linkage	Complete Linkage	Average Linkage	Centroid	Ward
Elbow	3	5, 9	4, 8	3, 6	6 (3)
R^2	(10)	7	6, 9, 7	7	(6)
R^2_{semi}	(5, 7, 10)	7	9, 6, 7	7	6
Pseudo F	5 (7, 10)	7	9, 11	7	(6)
Pseudo t^2	12 (3, 5)	9 (7)	7, 11	7 (5)	12, 4 (7)
RMSSTD	---	12, 8	(9)	---	10, 3, 6

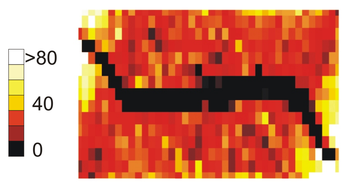
Estimated numbers of clusters for the standardized SEM-EEG data set

Measure	Method				
	Single Linkage	Complete Linkage	Average Linkage	Centroid	Ward
Elbow	3 (7, 11)	3 (9)	12 (6, 9)	6 (5)	4 (7, 9)
R^2	---	8, 4	(10)	7	3 (4)
R^2_{semi}	9, 11 (6)	8, 4	---	7	3, 4
Pseudo F	---	8, 4	4 (10)	7	(3)
Pseudo t^2	11 (9, 6, 3)	9, 5	8, 10	(11, 9, 5)	(10, 8)
RMSSTD	---	8	---	---	7, 12

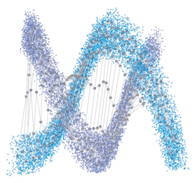
Estimated numbers of clusters for the non-standardized SEM-EEG data set

Winner Histogram Evaluation

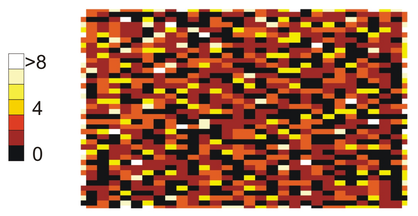
Winner Histogram for a Self-Organizing Feature Map



Artificial
Helix Data
20,000 Input Vectors
2 Clusters
3 dim. Input Vectors



Artificial
Gauss-Mixture Data
1652 Input Vectors
5 Clusters
47 dim. Input Vectors



Experimental
SEM-EEG Data
1652 Input Vectors
unknown Number of Clusters
47 dim. Input Vectors

U-Matrix Evaluation

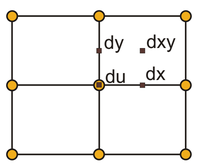
Unified Distance Matrix (U-Matrix) [Ultsch et al., 89]

Method for the Visualization of Distances between topological adjacent Prototype Vectors

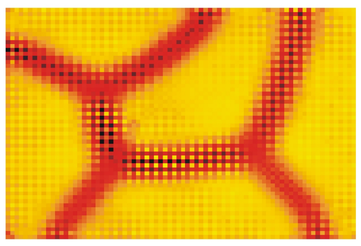
$$dxy(x,y) = \frac{1}{2}\left(\frac{\|\mathbf{w}_{x,y} - \mathbf{w}_{x+1,y+1}\|}{\sqrt{2}} + \frac{\|\mathbf{w}_{x,y+1} - \mathbf{w}_{x+1,y}\|}{\sqrt{2}}\right)$$

$$dy(x,y) = \|\mathbf{w}_{x,y} - \mathbf{w}_{x,y+1}\|$$

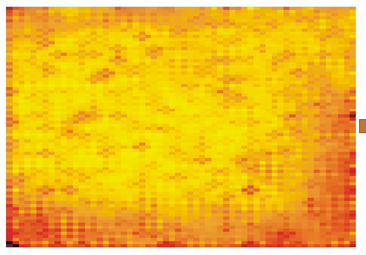
$$dx(x,y) = \|\mathbf{w}_{x,y} - \mathbf{w}_{x+1,y}\|$$



$$\mathbf{U} = \begin{bmatrix} du(1,1) & dx(1,1) & du(2,1) & \dots & du(n_x,1) \\ dy(1,1) & dxy(1,1) & dy(2,1) & \dots & dy(n_x,1) \\ du(1,2) & dx(1,2) & du(2,2) & \dots & du(n_x,2) \\ dy(1,2) & dxy(1,2) & dy(2,2) & \dots & dy(n_x,2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ du(1,n_y) & dx(1,n_y) & du(2,n_y) & \dots & du(n_x,n_y) \end{bmatrix}$$



U-Matrix for Gauss-Mixture Data

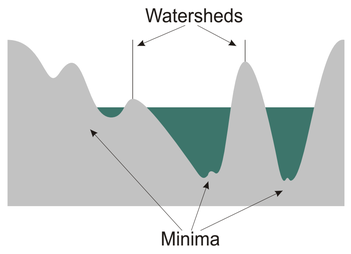


U-Matrix for SEM-EEG Data

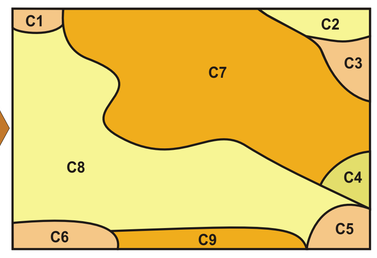
Segmentation

Segmentation with Watershed Transform [Beucher et al., 79]

Method for Image Segmentation



Steps:
1) Smooth the Function U(x,y)
2) Set initial Ground Level h_{min}
3) Mark all Minima Regions
4) Flood the Minima Regions
5) Install water-dividing Barriers



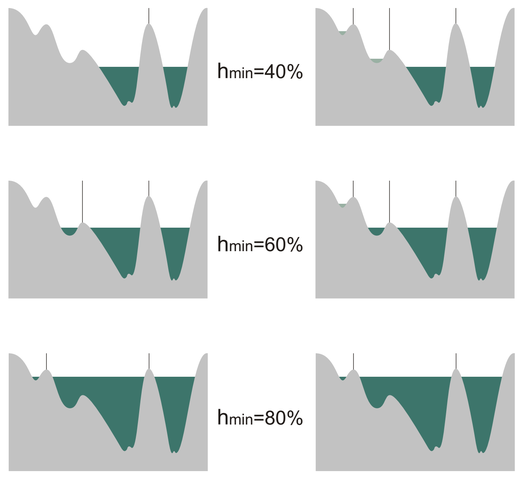
Segmented SOM of SEM-EEG Data
Cluster C1 to C9

Number of Clusters

Estimate the Number of Clusters [Costa et al., 99]

Method a: without Generation of new Minima

Method b: with Generation of new Minima



Results

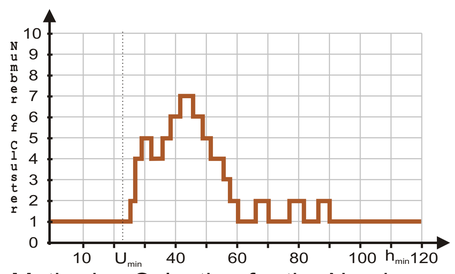
The "Segmented U-Matrix Method" is applied to the SEM-EEG Data Set in order to evaluate the Number of Clusters as a Function of the Ground Level (h_{min})

Watershed Transformation can be helpful, if Scoring of the U-Matrix is not evidently

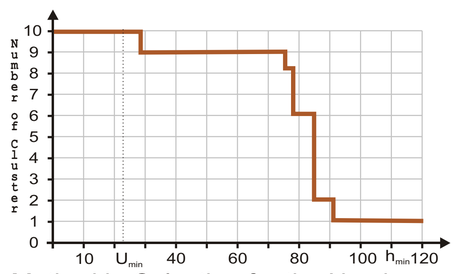
SEM-EEG is arrangeable in 9 Clusters

* three of them are of Alpha - Type (7.5 - 12.5 Hz)
* two of them are of Delta - Type (1.0 - 3,5 Hz)
* one of them is of Theta - Type (3.5 - 7.5 Hz)

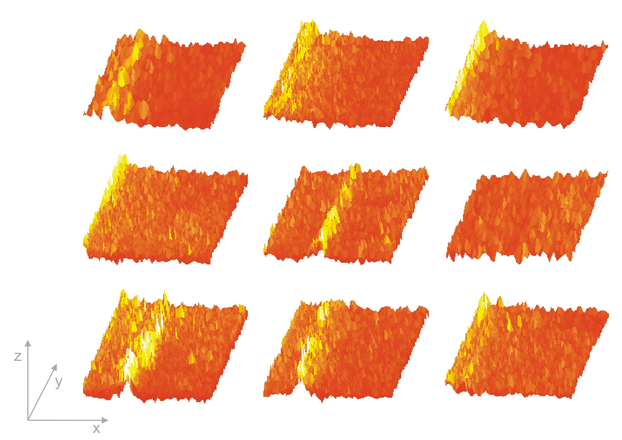
Limitation to two- dimensional Maps is not required



Method a: Selection for the Number of Clusters is difficult



Method b: Selection for the Number of Clusters is straight forward



Waterfall-Plot-Representation of the Cluster Solution
(x: Frequency, y: Item, z: Spectral Power Density)

University of Applied Sciences Schmalkalden
Department of Computer Science, PF 182
D-98574 Schmalkalden, Germany

fhS

{sommer, golz}@informatik.fh-schmalkalden.de
http://www.sund.de
http://www.sund.de/golz