Proceedings of TMCE 2012, May 7–11, 2012, Karlsruhe, Germany, Edited by I. Horváth, A. Albers, M. Behrendt and Z. Rusák

CLUSTERING DESIGNERS' MENTAL ACTIVITIES BASED ON EEG POWER

Thanh An Nguyen
Yong Zeng
Concordia Institute for Information Systems Engineering
Concordia University
{ngu_th, zeng}@encs.concordia.ca

ABSTRACT

To develop effective design methods and to improve design technologies, it is important to understand design cognitive activities. Following our previous preliminary work on the application of electrical brain signals to study designer behaviours, we further investigate the relationship between electrical brain signals and design activities. Firstly, we segmented design protocol based on design activities. Second, the segments were clustered through their corresponding brain signals. Third, we identified the similarity of the design activities among all of the segments belonging to the same cluster. It was observed that there was an association between certain types of design activities and brain's electrical signals. The result is promising and future research questions are identified to improve the proposed approach

KEYWORDS

Design cognition, clustering, electroencephalography (EEG), electrical brain signal, protocol analysis

1. INTRODUCTION

In order to develop effective design method and to improve current design technology, it is necessary to understand the natural design process [1]. Researchers in design have been trying different techniques to reveal the cognitive process underlying the observed designer behaviours. The most popular technique adopted in design research is verbal protocol analysis. Verbal protocol analysis dates back to Aristolte's time [2]. In design research, protocol analysis was first adopted in late 1960s [3] and has increased quickly after 1994 [4]. In particular, researchers have used protocol analysis to study the differences in design approach between novice and experienced designers [5-7], to investigate the impact of language on design activities [8], to study the decision making in design [9], and to validate design model [10,11]. The relationship between creativity and design process is also examined in different aspects [12,13]. The disadvantage of protocol analysis method is that researchers can only study behaviours that are observable or behaviours that can be verbalized. Therefore, it is necessary to find a new method which can help researchers study complete cognitive process.

Researchers in brain-computer interface (BCI), Kansei (emotional) engineering, human factors science and psychology have been focusing on using physiological signals to control external devices, identify human's emotional state, explore human capability [14-16] and study mental activities [17-19]. Following the trend, researchers start to use physiological signals to study design activities [20,21]. The method allows researchers to directly observe changes in subject's physiological signals which may reflect changes in emotion and/or cognitive function.

Continuing our previous preliminary work on the application of electrical brain (EEG) signals to study designer behaviors [22], we further investigate the relationship between EEG signals and design activities. In the current study, we group design activities based on the corresponding electrical brain signals recorded during the design process.

The paper is organized as follows: a brief introduction to EEG is given in Section 2 followed by an introduction to clustering k-means algorithm in Section 3. Section 4 and Section 5 present experimental setting and data analysis. Section 6 presents the result and Section 7 concludes the paper.

2. ELECTROENCEPHALOGRAPHY (EEG)

Electroencephalography (EEG) is the method of recording brain's electrical signals. The EEG picks up the signals originated from the communication between neurons, which is the process of releasing neurotransmitters from one neuron (pre-synaptic neuron) to adjacent neuron (post-synaptic neuron)

[23]. The release of neurotransmitters to the post-synaptic neuron causes chemical ions to move between the extracellular and intracellular environment. This movement leads to the change in extracellular and intracellular potential and make the neurons become a dipole. Thousands of such dipoles will generate a voltage large enough to propagate to the scalp and to be picked up by EEG system [24].

Electrical brain signals, also called EEG signals, are classified in terms of frequency band. There are four main bands: delta (1-4Hz), theta (4-8Hz), alpha (8-13Hz) and beta (>13Hz). The delta rhythm is present all over the scalp during sleeping state. Theta rhythm is dominant during relaxed state and eye open. Alpha rhythm is dominant during eye close and beta rhythm appears when the brain is engaged in visual or cognitive activities [25].

The position of EEG sensors usually follows international standard 10/20, 10/10 or 10/5 systems. In the current paper, EEG is recorded based on 10/20 system. The 10/20 system divides a head into 19 locations. Each location is denoted by two parts; the first part is letters indicating the lobe of the brain: F is for frontal lobe, Fp is for prefrontal cortex (part of frontal lobe), P is for parietal lobe, O is for occipital lobe, T is for temporal lobe and C is for motor cortex (part of frontal lobe). The second part is a number or a letter z: odd number indicating left hemisphere, even number indicating right hemisphere and z indicating the midline.

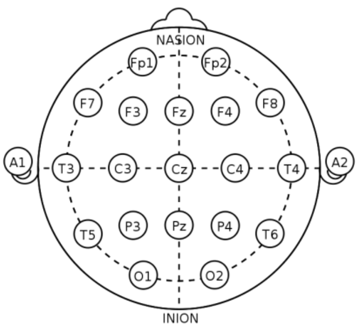


Figure 1 The 10/20 system of the brain [26]

3. CLUSTERING

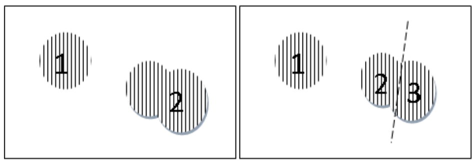
Clustering is the unsupervised classification of data into groups based on similarity [27]. Clustering of EEG signals has been used widely to identify phenomenon of interest [28-31]. In the current paper, k-means is adopted to cluster EEG signals. In k-means, the number of clusters has to be defined in advance. The objective of k-means algorithm is to minimize the cost function [32-34]

$$E = \sum_{l=1}^{k} \sum_{i=1}^{n} y_{il}\, d(X_i, Q_l) \qquad (1)$$

where k is the number of clusters, n is the number of objects, X_i is the object i and Q_l is the mean of cluster l, y_{il} is the weight associated with object i in relation to cluster l, $d(X_i, Q_l)$ is the similarity measure between object X_i and cluster mean Q_l,

The algorithm k-means randomly assigns each data object in a dataset to arbitrary clusters. The mean of each cluster is calculated and each data object is evaluated against the cluster mean and is re-allocated if necessary. The process continues until no reallocation occurs.

Most of the clustering algorithms require us to pre-specify certain parameters. If the parameters are not chosen properly, the resulting clusters will not be optimal. An example of an optimal and a non-optimal clustering result is given in Figure 2(a). Visually, the optimal number of clusters should be two as depicted in Figure 2(a) but an improper parameter assigned to the algorithm can result in three clusters as shown in Figure 2(b).



(a) (b)

Figure 2 Optimal and non-optimal number of clusters

Therefore, validation of the clustering results is necessary to obtain the optimal clusters. There exist several validation methods, a review of which is reported in [35]. The index used in the current study is Hartigan index.

The Hartigan index [36] is defined as follows:

$$H^k = (n - k - 1)\left(\frac{SS_k}{SS_{k+1}} - 1\right) \qquad (2)$$

where H^k is the Hartigan index of k clusters, n is the total number of objects in a data set, SS_k is the sum of square of distance from the objects to their cluster centroids when there is k clusters and SS_{k+1} is the sum of square of distance from the objects to their cluster centroids when there is (k+1) clusters. According to Hartigan, the optimal number of clusters is the smallest k such that $H^k \leq 10$.

4. EXPERIMENT

4.1. Participants

There are totally 13 participants from different ethnic background with age ranging from 20 to 40 years old. All of the participants are graduate students from engineering department and have English as either foreign language or first language. All of the participants have never taken any specific design training and are not aware of any design methods. In the current study, analysis for one subject is reported.

4.2. Design problem

Subjects are asked to work on the following design problem:

"Design a house that can easily fly from one place to another place.

*In your final result, clearly specify the **functions** (what can it do? how...) and the **features** of your product (How does it look like...)"*

This type of problem is chosen to minimize the domain-dependent effect.

4.3. Experiment Process

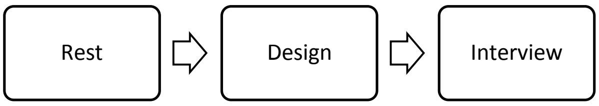


Figure 3 Experiment process: designer rests for 6 six minutes, works on the design problem and eventually are interviewed by the experimenter

The experiment is divided into three main parts as shown in Figure 3. The first part lasted for six minutes. In the first three minutes the subjects are asked to relax with eye opened, in the next three minutes the subjects are asked to relax with eye closed. The second part is design experiment. The subjects are asked to work on a given design problem. There is no time limit. The third part is the interview process. After completing the design, the subjects are interviewed to report what they were thinking of during the design task. The recorded video is shown to the designers as a memory cue to answer the interviewed questions.

4.4. Experiment Setup

The devices used in the experiment are: camera, heart rate variability (HRV), tablet and EEG system. Details for each device are listed below:

1. Camera: There are four main cameras recoding designers: one for the upper body part, one for the lower body part, one for the whole body and one for the face.
2. HRV recorder: We use polar RS8000 device to record the variation in heart beat intervals.
3. Tablet: the subject use tablet to do the design. Screen recording software records all the design activities.
4. EEG system: We record 14 channels: Fpz, Fz, F4, F3, C4, C3, T4, T3, T6, T5, P4, P3, O2, O1 and right ear as shown in Figure 4. All the recordings are referential to left ear. The signal is then reconstructed into digital linked-ear montage.

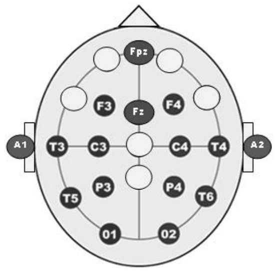
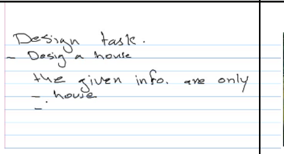
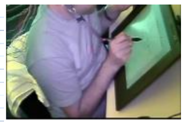
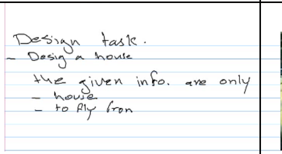
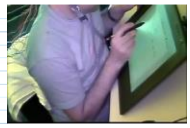
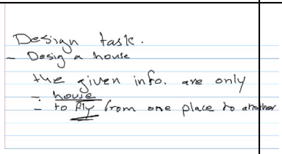
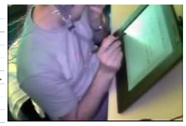


Figure 4 Fourteen channels and right earlobe A2 recorded by EEG

The experiment setting is shown in Figure 5.

Figure 5 Experiment setting

Table 1 Examples of video segments

Segment	Start	End	Action observed from the screen	Designer behaviors	Video data	Video data
8	11:25.391	11:33.591	No action	Lift pen from the tablet, look at the screen		
9	11:33.591	11:53.351	Write	Write, left hand put on the side of the tablet		
10	11:53.351	11:59.561	No action	Scratch eyes and look up for a few seconds		

5. DATA PROCESSING AND ANALYSIS

5.1. Data segmentation

The tablet screen recording video serves as the basis for the segmentation. The video is segmented based on designer behaviors. The discontinuity in designer's actions marks the starting and ending of a segment as illustrated in Figure 6. Examples of video segments are shown in Table 1.

5.2. Data coding

Data was coded based on the following pre-defined design activities: search for knowledge, analyze knowledge, evaluate knowledge, identify requirements, analyze requirements, express requirements, generate solution, express solution, and evaluate solution.

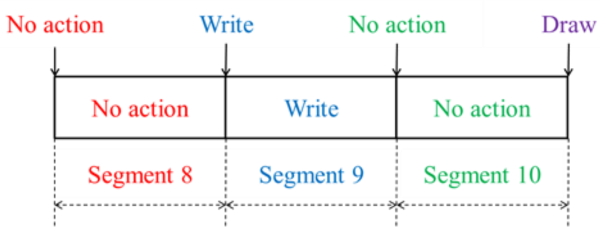


Figure 6 An example of data segments

The data is coded by two researchers. Any discrepancies in the coded data were discussed and resolved.

5.3. EEG data processing and analysis

The raw EEG data was passed through a high pass filter 0.3Hz, a low pass filter 40Hz and a Notch filter 60Hz. Artifact caused by excessive movement were manually detected and removed. EEG data were segmented according to the video segments. Originally there were 54 segments in total. After the artifact removal step, there were 34 segments left.

Muscular artifact and ocular artifact is removed from each EEG segment by EEGLab plug-in AAR toolbox [37,38].

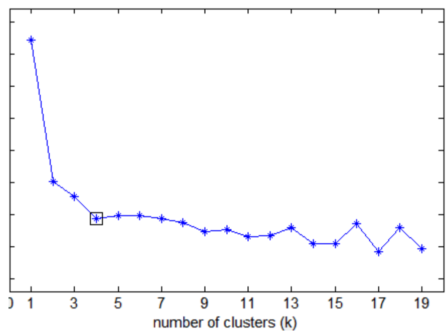


Figure 7 Hartigan index shows that the optimal number of clusters is 4

The differences in power spectral density between the right and the left hemisphere were computed for each EEG segment. The result was used as features for k-means clustering algorithm. Validity index was calculated to obtain the optimal number of clusters. The clustering and cluster validation were computed by CVAP toolbox [39]. For the current data set, Hartigan index shows four as the optimal number of clusters as presented in Figure 7.

6. RESULTS

The 34 segments are grouped into 4 clusters. Table 2 shows the number of segments for each cluster. Table 3 shows the percentage of segments distributed over the design activities and the corresponding cluster number. A graphical result is shown in Figure 8.

Table 2 Distribution of clusters

Cluster number	Number of segments
1	1
2	4
3	13
4	16
Total	34

Table 3 Clustering result

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Id-R, Ev-R	0	0	0	2.94%
An-R, Ev-R	0	0	2.94%	0
Sr-K	0	0	2.94%	5.88%
Sr-K, Gr-S	0	0	2.94%	2.94%
Sr-K, An-K, Ev-K	0	2.94%	0	0
Sr-K, Ev-S, Gr-S, Ex-S	2.94%	0	0	0
An-K, Ev-K, Gr-S	0	0	0	2.94%
Ex-S	0	5.88%	14.7%	20.59%
Ex-S, Ev-S	0	0	2.94%	0
Ev-S, Gr-S	0	0	5.88%	2.94%
Ex-S, Gr-S	0	0	2.94%	0
Ev-S	0	0	0	2.94%

Sr-K: search for knowledge
An-K: analyze knowledge
Ev-K: evaluate knowledge
Ex-S: express solution
Ev-S: evaluate solution
Gr-S: generate solution
Ev-R: evaluate requirement
Id-R: identify requirement
An-R: analyze requirement

From Table 3 and Figure 8, we see that design activities belong to cluster 3 are: {solution expression, solution generation}, {solution evaluation, solution generation}, {solution expression, solution evaluation}, {solution expression}, {knowledge search, solution generation}, {knowledge search}, {requirement analysis, requirement evaluation}.

Solution generation and evaluation activity appear most frequently in cluster 3. The reason that some segments are not coded as solution generation and/or evaluation can be that it is not always clear for researchers to identify solution generation and/or evaluation for a certain segment. Idea generation and evaluation can be embedded in other activities such as solution expression, knowledge search, requirement analysis, and requirement evaluation.

Furthermore, the second last segment, in which designers are about to finish the design and solution generation is unlikely to occur, belongs to cluster 3. This makes us believe that evaluation plays a more dominant role in cluster 3 than solution generation. Therefore, it is possible that most of the design activities in cluster 3 include evaluation activity.

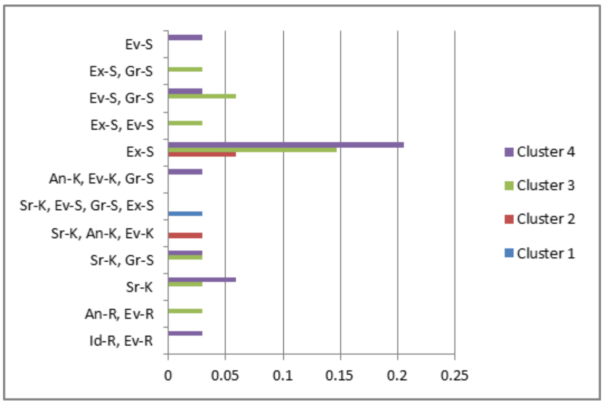


Figure 8 Clustering result

Design activities belong to cluster 4 are: {solution **evaluation**}, {solution expression, solution generation}, {solution expression}, {knowledge analysis, knowledge **evaluation**, solution generation}, {knowledge search, solution generation}, {knowledge search}, {requirement identification, requirement **evaluation**}.

Similarly, evaluation and generation appear more frequently in cluster 4. Based on the protocol, after reading the design problem, the designer comes up with a solution first before analyzing the requirement. Cluster 4 appears more frequent in the beginning of the design process and less in the end of the design process. Therefore, cluster 4 should represent solution generation.

Activities belong to cluster 1 and cluster 2 are unclear due to small number of segments for the clusters. However, cluster 2 tends to be a combination of several design activities.

It is noted that although most members of cluster 4 and cluster 3 belong to solution expression, neither cluster 3 nor cluster 4 can be taken as solution expression because solution expression is not a primitive activity. Expression often includes evaluation and/or generation activities.

7. CONCLUDING REMARKS AND DISCUSSION

In the paper, we cluster the segments based on the differences in EEG power spectral density between the right and left hemisphere. The data set is from one subject. The purpose of the analysis is to find potential approach to studying design activities using EEG signals rather than address the issue of generality.

Our hypothesis is that similar activities observed from the segments should have similar EEG features and appropriate clustering method should be able to group similar-activity segments based on EEG features. In conducting the study, we adopted k-means algorithm to cluster the segments. We found that solution generation and evaluation are likely to intertwine but the role of one can be more dominant than the role of the other in certain segments. Solution expression and knowledge search are not primitive design activities. For future work, experiments will be done to find primitive cognitive design activities. Segmentation scheme will be revised to better fit the EEG-based analysis approach.

AKNOWLEDGEMENT

This research is partially funded by Natural Sciences and Engineering Research Council of Canada (NSERC) through a Discovery Grant. The first author gratefully thanks NSERC for the financial support through a Canada Graduate Scholarship (CGS).

REFERENCES

[1] Zhu, S., Yao, S., and Zeng, Y., (2007), "A novel approach to quantifying designer's mental stress in the conceptual design process", ASME Conference Proceedings, Vol. 2007 (48035), pp. 593-600.

[2] Ericsson, K.A., and Crutcher, R.J., (1991), "Introspection and verbal reports on cognitive processes--Two approaches to the study of thinking: A response to Howe", New Ideas in Psychology, Vol. 9 (1), pp. 57-71.

[3] Eastman, C.M., (1969), "Cognitive processes and ill-defined problems: a case study from design", Proceedings of the Proceedings of the 1st international joint conference on Artificial intelligence, Morgan Kaufmann Publishers Inc., Washington, DC.

[4] Jiang, H., and Yen, C.C., (2009), "Protocol analysis in design research: a review", Proceedings of the International Association of Societies of Design Research (IASDR) Conference, Seoul, Korea.

[5] Ahmed, S., and Christensen, B.T., (2008), "Use of analogies by novice and experienced design engineers", ASME Conference Proceedings, Vol. 2008 (43284), pp. 11-19.

[6] Atman, C.J., Chimka, J.R., Bursic, K.M., and Nachtmann, H.L., (1999), "A comparison of freshman and senior engineering design processes", Design Studies, Vol. 20 (2), pp. 131-152.

[7] Suwa, M., and Tversky, B., (1997), "What do architects and students perceive in their design sketches? A protocol analysis", Design Studies, Vol. 18 (4), pp. 385-403.

[8] Dong, A., Davies, K., and McInnes, D., (2005), "Exploring the Relationship Between Lexical Behavior and Concept Formation in Design Conversations", Proceedings of the ASME Conference Proceedings, Vol. 2005, California, USA, pp. 61-71.

[9] Akin, Ö., and Lin, C., (1995), "Design protocol data and novel design decisions", Design Studies, Vol. 16 (2), pp. 211-236.

[10] Neill, T.M., Gero, J.S., and Warren, J., (1999), "Understanding Conceptual Electronic Design Using Protocol Analysis", Research in Engineering Design, Vol. 10 (3), pp. 129-140.

[11] Jin, Y., and Benami, O., (2010), "Creative patterns and stimulation in conceptual design", Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Vol. 24 (2), pp. 191-209

[12] Kim, Y.S., Jin, S.T., and Lee, S.W., (2006), "Design activities and personal creativity characteristics: A Case Study of Dual Protocol Analysis Using Design Information and Process", ASME Conference Proceedings, Vol. 2006 (42584a), pp. 181-190.

[13] Badke-Schaub, P., Goldschmidt, G., and Meijer, M., (2010), "How Does Cognitive Conflict in Design Teams Support the Development of Creative Ideas?" Creativity and Innovation Management, Vol. 19 (2), pp. 119-133.

[14] Nagashima, T., Tanaka, H., and Uozumi, T., (2008), "An overview of Kansei engineering: a proposal of Kansei informatics toward realising safety and pleasantness of individuals in information network society", Int. J. Biometrics, Vol. 1 (1), pp. 3-19.

[15] Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., McFarland, D.J., Peckham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J., and Vaughan, T.M., (2000), "Brain-computer interface technology: a review of the first international meeting", Rehabilitation Engineering, IEEE Transactions on, Vol. 8 (2), pp. 164-173.

[16] Lambooij, M., IJsselsteijn, W., Fortuin, M., and Heynderickx, I., (2009), "Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review", Journal of Imaging Science and Technology, Vol. 53 (3), pp. 1-14.

[17] Mihov, K.M., Denzler, M., and Förster, J., (2010), "Hemispheric specialization and creative thinking: A meta-analytic review of lateralization of creativity", Brain and Cognition, Vol. 72 (3), pp. 442-448.

[18] Khader, P.H., Jost, K., Ranganath, C., and Rösler, F., (2009), "Theta and alpha oscillations during working-memory maintenance predict successful long-term memory encoding", Neuroscience Letters, Vol. 468 (3), pp. 339-343.

[19] Sobolewski, A., Holt, E., Kublik, E., and Wróbel, A., (2011), "Impact of meditation on emotional processing - A visual ERP study", Neuroscience Research, Vol. 71 (1), pp. 44-48.

[20] Göker, M.H., (1997), "The effects of experience during design problem solving", Design Studies, Vol. 18 (4), pp. 405-426.

[21] Alexiou, K., Zamenopoulos, T., Johnson, J.H., and Gilbert, S.J., (2009), "Exploring the neurological basis of design cognition using brain imaging: some preliminary results", Design Studies, Vol. 30 (6), pp. 623-647.

[22] Nguyen, T.A., and Zeng, Y., (2010), "Analysis of design activities using EEG signals", ASME Conference Proceedings, Vol. 2010 (44137), pp. 277-286.

[23] Kirschstein, T., and Köhling, R., (2009), "What is the Source of the EEG?" Clinical EEG and neuroscience, Vol. 40 (3), pp. 146-149

[24] Aguiar, P., David, A., Paulo, S., and Rosa, A., (2000), "EEGSOLVER — brain activity and genetic algorithms", Proceedings of the 2000 ACM symposium on Applied computing - Volume 1, 2000.

[25] Van De Graaff, K.M., (2001), Human anatomy, 6 ed., McGraw Hill Higher Education.

[26] Wikipedia, (2012), "10-20 system (EEG)", Wikipedia. The free Encyclopedia 2012.

[27] Jain, A.K., Murty, M.N., and Flynn, P.J., (1999), "Data clustering: a review", ACM Comput. Surv., Vol. 31 (3), pp. 264-323.

[28] [Alkan, A., and Akben, S.B., (2011), "Use of K-means clustering in migraine detection by using EEG records under flash stimulation", International Journal of the Physical Science, Vol. 6 (4), pp. 641-650.

[29] Geva, A.B., and Kerem, D.H., (1998), "Forecasting generalized epileptic seizures from the EEG signal by wavelet analysis and dynamic unsupervised fuzzy clustering", Biomedical Engineering, IEEE Transactions on, Vol. 45 (10), pp. 1205-1216.

[30] Murugappan, M., Rizon, M., Nagarajan, R., and Yaacob, S., (2009), "FCM clustering of Human Emotions using Wavelet based Features from EEG", Biomedical Soft Computing and Human Sciences, Vol. 14 (2), pp. 35-40.

[31] Bankman, I., and Gath, I., (1987), "Feature extraction and clustering of EEG during anaesthesia", Medical and Biological Engineering and Computing, Vol. 25 (4), pp. 474-477.

[32] Lloyd, S.P., (1952), "Least square quantization in PCM", Bell Telephone Laboratories, 1952.

[33] MacQueen, J., (1967), "Some methods for classification and analysis of multivariate observations", Proceedings of the Proc. Fifth Berkeley Symp. on Math. Statist. and Prob, Vol. 1, Univ. of Calif. Press, pp. 281-297.

[34] Suh, S.C., (2010), Practical applications of data mining, Jones & Bartlett Learning.

[35] Halkidi, M., Batistakis, Y., and Vazirgiannis, M., (2002), "Clustering validity checking methods: part II", SIGMOD Rec., Vol. 31 (3), pp. 19-27.

[36] Hartigan, J.A., (1975), Clustering algorithms, Wiley.

[37] Gomez-Herrero, G., De Clercq, W., Anwar, H., Kara, O., Egiazarian, K., Van Huffel, S., and Van Paesschen, W., (2006), "Automatic Removal of Ocular Artifacts in the EEG without an EOG Reference Channel", Proceedings of the Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic, pp. 130-133.

[38] Gómez-Herrero, G., (2007), "Automatic Artifact Removal (AAR) toolbox v1.3 (Release 09.12.2007) for MATLAB", Vol. 2010, Tampere University of Technology, 2007.

[39] Wang, K., Wang, B., and Peng, L., (2009), "CVAP: validation for cluster analyses", Data Science Journal, Vol. 8, pp. 88-93.