Clustering of advertising images using electroencephalogram

Ingon Chanpornpakdi*, Motoi Noda*, Toshihisa Tanaka*, Yuval Harpaz‡, and Amir B. Geva†‡
* Tokyo University of Agriculture and Technology, Tokyo, Japan
E-mail: mint@sip.tuat.ac.jp, s203382s@st.go.tuat.ac.jp, tanakat@cc.tuat.ac.jp Tel/Fax: +81-42-388-7123
† Ben-Gurion University of the Negev, Beersheba, Israel
E-mail: geva@ee.bgu.ac.il Tel/Fax: +972-8-646-1600
‡ InnerEye Ltd, Herzliya, Israel
E-mail: yuval@innereye.ai Tel: +972-9-749-0854

Abstract—**Packaging and advertisements of brands affect customers' decision-making on purchasing products and could lead to business loss. Hence, neuromarketing, the application of neuroscience in the marketing field, is introduced aiming to understand customers' cognitive functions toward advertisements or products. Our study focused on identifying how the brain respond to different types of advertising image of the same brand were perceived using electroencephalogram (EEG). We performed an experiment using 33 different Coca-Cola advertising images in RSVP (rapid serial visual presentation) task on 23 participants. A seven channels EEG dry headset was used to record the visual event-related potential (ERP), specifically, the positive peak found at 300 to 700 ms after image onset; P300, to compare the perception response. We applied k-means and hierarchical clustering to the obtained EEG data, and achieved the best clustering for three clusters, yielding different P300 amplitudes and latencies. The typical Coca-Cola ads, red color with Cola-cola text on the ads, induced a faster and larger response, implying better perception than the unconventional or black color ads. We conclude that ERP clustering may be a useful tool for neuromarketing. However, the relationship between the EEG-based cluster and the image-based cluster should be further investigated to confirm the suggestion.**

I. Introduction

Many companies rebrand existing products to promote their sales. However, rebranding is a risky action since the customers may fail to identify brands strongly associated with specific image. One of the famous rebranding case studies is Tropicana's repackaging failure.

In 2009, Tropicana redesigned their package from a classical straw-in-orange package into a clean and simple package with the close-up look of a glass of orange juice. They wanted to stress that their orange juice was 100% pure, natural, and freshly squeezed. Unfortunately, this rebranding failed because the brand lost its identity and attracted less attention on the shelves with the new package. According to the customers, the new design turned out to resemble to generic low-cost brands, and the font and color of "Tropicana", which had been important for identifying the product, was disfigured. This failure costed approximately $100 million sales loss in two months and $35 million in package redesigning project. [1], [2]

The aforementioned study showed that most influencing factor in rebranding is the customer perception [3]. Traditionally, in order to have an insight into customers' preferences, simple questionnaire studies are executed. However, such methods may lead to inaccurate information, due to biases depending on the question asked and the criteria of participants [4]. For that reason, neuromarketing; the application of neurotechnologies, sometimes combined with biosignal changes, to measure the brain response in the marketing field, has been receiving a lot of attention from researchers. With this method, unbiased responses are directly decoded from customers' brains [5], [6].

The neuroimaging technologies used to identify customers' preferences were electroencephalography (EEG), functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and transcranial magnetic stimulation (TMS) [7]. In addition to the brain response, other biosignals, typically eye-tracking and galvanic skin response, are frequently recorded to obtain information. The stimulus of many sensory modalities; visual, auditory, or the combination of both visual and auditory stimuli, were often utilized in neuromarketing research [8]. In this study, we used EEG in our investigation of brain response on different advertising images, due to its simplicity, portability, low cost, and high temporal resolution [9].

A common method for exploring visual cognition using EEG is by displaying a stream of images at a rapid rate, often multiple images per second. This presentation is called rapid serial visual presentation (RSVP) [10]. The brain response corresponding to the displayed images is called event-related potentials (ERP), typically named by their polarity and latency evoked after the stimulus was shown [11]. It has been reported that N200, negative potentials evoked around 200 msec post-stimulus, was related to the recommendation of images in advertising image cognition whereas P300, positive peak evoked around 300 to 700 msec post-stimulus, was related to the advertising image cognition with high familiarity in online shopping recommendation study [12]. However, the criteria of advertising image that could enhance image perception of the customer are still unknown.

As neuromarketing is designed to be used outside the laboratory, we focused on the real-life application using the dry

EEG headset, which has high feasibility and requires less time in preparation when compared to the wet EEG headset. Our aim of this study is to understand what types of advertisements could enhance the cognition of the brand, and to suggest a method for evaluating how effective advertising images are in drawing attention on social media or even on the shelves.

II. Material and Methods

A. Participants

Twenty-three volunteers (thirteen male and ten female with an average age of 25.83 ± 4.03 and age range of 20 to 37) participated in the experiment. The volunteers were recruited from university students and working adults who do not have neurological disorders and self-reported normal or corrected-to-normal vision. None of the student-participants were encouraged to participate in this experiment by their professors, nor did they obtain any credits for doing so. Informed consent forms, approved by the Research Ethics Committee of the Tokyo University of Agriculture and Technology (N02-14-E95), were obtained from all participants before the experiment.

B. Experimental Design

The stimuli were created from 266 advertising images (33 Coca-Cola advertising images as targets and 233 unique advertising images as non-targets). Each image was repeated six times and randomly presented at 3 Hz using EEG recording software version 0.4.27 home-tailored by InnerEye Ltd. shown in Fig. 1. The order of images in stimulus presentation was set so that there were no two target images presented consecutively. The task was divided into 11 one-minute blocks and allowed the participants to rest after each block ended. The participants were instructed to count the number of times that the Coca-Cola advertising images appeared, and were allowed to check their count performance by comparing the correct number of times that the target image shown provided at the end of each block. They were also asked to report their impression on how correctly they could recognize the images.

C. Data Acquisition

We recorded EEG at 300 Hz sampling frequency using seven channels DSI-7 dry headset (DSI-7, Wearable Sensing, USA). The electrode montage of the DSI-7 headset was F3, F4, C3, C4, Pz, PO7, and PO8 positions according to the 10-10 international system, with the ground electrode at the Fpz position (Fig. 2). We used the average input of A1 and A2 electrodes around the left and right earlobe as a reference. The participants sat at a distance of approximately 60 cm away from the monitor shown in Fig. 3. The participants were asked to keep their eyes on the screen to avoid getting distracted by the environment of the experimental room.

D. EEG Pre-processing

The EEG analysis was processed using MNE- python package 1.3.0 [13] on Python 3.5.8. The recorded EEG was filtered with 1 to 20 Hz FIR bandpass filter [14] and down-sampling

Fig. 1. Cola-cola advertisement image stream of RSVP stimulus used in the experiment

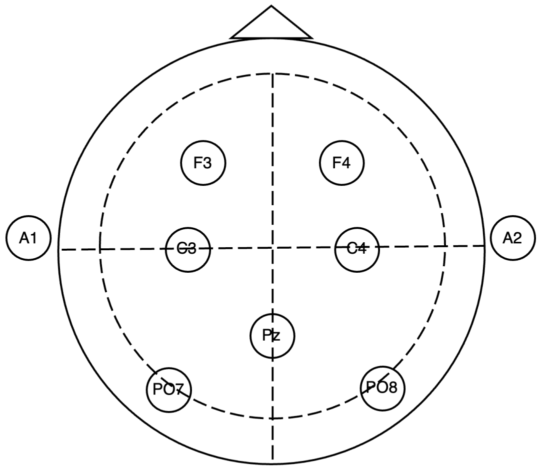


Fig. 2. Electrode positions

to 150 Hz. The eye blink artifact was removed using fastICA method of independent component analysis (ICA) algorithm. Seven independent components (ICs) were obtained and 1.57 ± 0.59 ICs were removed with the removal range of one to three ICs in eye blink removal process. The obtained EEG was epoched into 1,596 epochs (198 target and 1,398 non-target epochs) from -0.1 sec to 1.0 sec and the epochs with amplitude higher than 100 µV were rejected to remove the noise contaminated epochs. The EEG epochs of each Coca-Cola advertising image (33 images in total) were averaged at each electrode position and grand averaged across the participants to calculate the ERP.

E. Clustering

The EEG epochs of each Coca-Cola advertising image (33 images in total) were averaged at each electrode position and grand averaged across the participants. The electrode positions were concatenated in the following order; F3, F4, C3, C4, PO7, PO8 and Pz. The grand averaged or per-image ERP data was

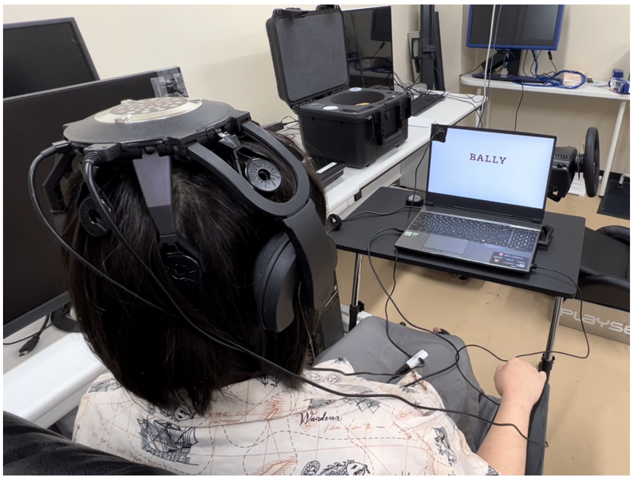


Fig. 3. The actual condition of the experiment

clustered using K-means clustering and hierarchical clustering algorithms. The duration of input EEG data was extracted at P300 duration of 0.3 sec to 0.7 sec. Hence, the dimension of input data was [33, 420]; the EEG at seven electrode positions containing 400 msec data for the perception of 33 images at 150 Hz sampling rate. To determine the best number of clusters, we implemented the elbow and silhouette method for K-means clustering and increment algorithm for hierarchical clustering models.

F. Statistical test

We performed a dependent t-test on testing the null hypothesis that there were no differences between P300 components when perceived target Coca-Cola images and non-target ads. After averaging all target and non-target images, the dependent t-test was executed using three features extracted from 0.40 to 0.70 sec duration; the average of all amplitudes during the extracted period, the maximum peak of P300, and latency of P300. We have performed a t-test on eight electrode conditions; averaged across all electrodes, and seven single electrode conditions.

III. Results

A. Event-Related Potential (ERP)

Fig. 4 shows the ERP obtained after grandaveraging the data of 23 participants. In this figure, we could observed relatively higher peaks between 0.30 to 0.70 sec after target onset, compared to non-targets. Since there were three images presented in one sec (presentation rate at 3 Hz), each trace contains the ERPs of three images. This is more visible for non-targets which do not have the P300 on top of the early response for all three images. The P300 shape contained two small peaks and found in all electrodes. The amplitude of the peaks were largest at Pz, PO7 and PO8 electrodes which located around the parietal and parietal-occipital area and smallest at the frontal F3 and F4.

The p-value obtained when compared the amplitude of P300 found during target and non-target presentation using three features; average amplitudes, maximum P300 amplitude, and latency of P300 are shown in Table I. The difference in latency of P300 during target and non-target image presentation was also calculated. The negative sign in latency difference denotes that the latency of P300 during non-target image presentation was longer than that of target image presentation. However, there was a significant difference between the latency of P300 evoked by target and non-target images on average across electrodes condition and only at F3, F4, C3, C4, and Pz electrode of the single electrode position condition The P300 latency of non-target images was significantly delayed at around 0.05 sec.

B. Clustering

To identify the types of image that caused two peaks of P300, we have performed target advertising image clustering using K-means and hierarchical clustering based on the P300 (duration of 300 to 700 msec). Fig. 5 and 6 show the number of clusters (k-value) optimization in K-means and hierarchical clustering, respectively. We observed that the best number of clusters in K-means clustering somehow did not yield a clear solution. In Fig. 5 (a), the result from elbow method shows that when $k = 6$, within cluster sum of squares calculated were decreasing linearly and the best k-value could be either 2, 3, or 6 clusters from human visualization. However, the result from silhouette method showed that 2 and 4 clusters were the most appropriate k-value. To clarify the best k-value, we have calculated the ERP of each group obtained after clustering for all the possibilities and found that when $k = 3$, the ERP shape of each group were able to differentiate with the least complexity, as shown in Fig. 7. In hierarchical clustering, we have performed k-value optimization using increment algorithm and found that the best cluster number was three clusters, as shown in Fig. 6. The ERP of three clusters obtained from the hierarchical clustering were plotted in Fig. 8.

From Fig. 7 and 8, the two clustering methods yielded similar results for the best three clusters. Each cluster has three different shapes of P300; largest amplitude at early latency, intermediate amplitude of two peaks at both early and late latency, and largest amplitude at late latency. The images clustered in each group were also similar in both K-means and hierarchical clustering methods. The P300 amplitude of when the participants perceived target and non-target advertising images were significantly different regardless of electrode positions; average and single electrode. The topography drawn in both Fig 7 and 8 show the source position of peaks at 0.40 and 0.50 sec respectively, with dark blue represents the minimum EEG amplitude and dark red represents the maximum EEG amplitude. We could see that the maximum amplitude of the ERP were at Pz position in all the group. Furthermore, we observed that group 0 in K-means and group 2 in hierarchical clustering having maximum amplitude around 0.40 s and identical amplitude at 0.50 s as well, unlike group 1 in both clustering which had noticeable peak only at 0.50 s.

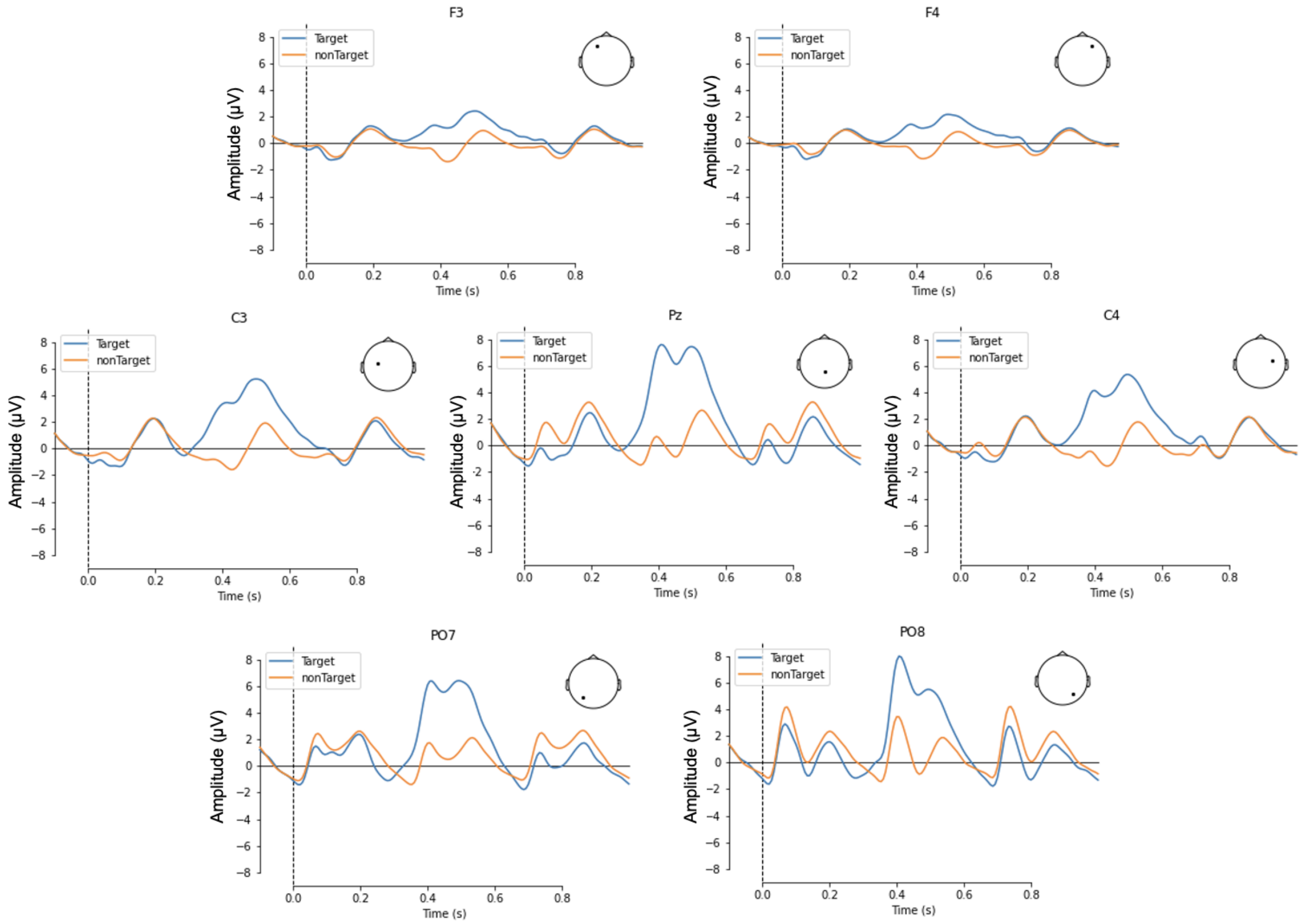


Fig. 4. ERP observed when grandaveraged across 23 participants where blue line and orange line represent the ERP evoked when the target and non-target images were perceived, respectively. The time zero was set to be the stimulus onset time

IV. Discussion

The P300 evoked for Coca-Cola advertising images cognition task were found to be larger than the other advertising images between 0.30 to 0.70 sec with two sub-peaks. When we clustered the the Coca-Cola images based on the evoked P300, we obtained the best result with three clusters of Coca-Cola advertising images which were able to disassemble the two sub-peaks of the P300.

A. Event-Related Potential (ERP)

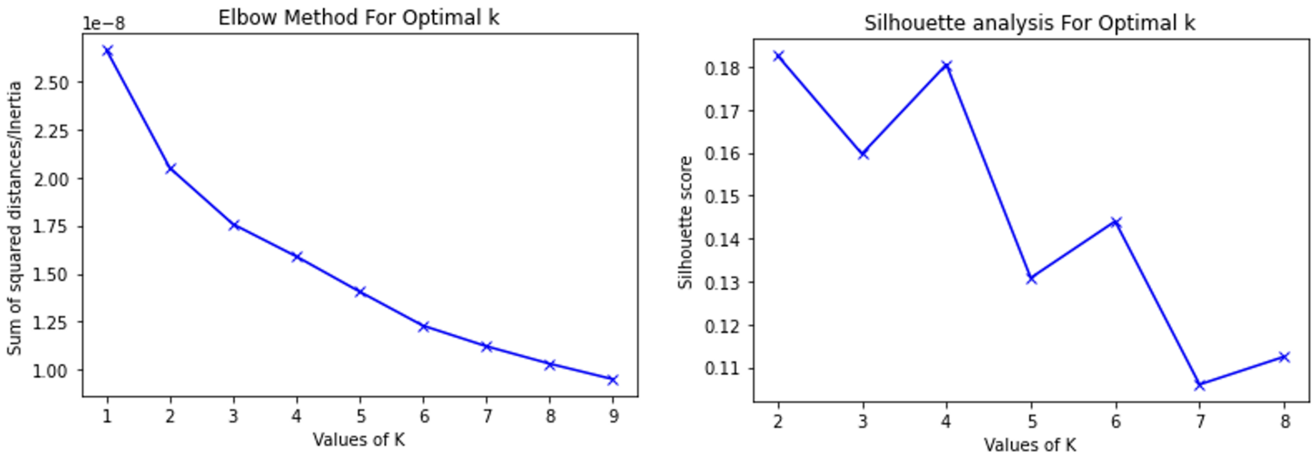
From Fig. 4, we could observe that there might be two sub-component of P300 potentials hidden in the ERP. Since Coca-Cola is the soft drink brand that all participants are familiar with, the observation of P300 was in line with the previous study, reported that advertising image with high familiarity evoked higher potential of P300 [12], that we observed two sub-components. In fact, P300 has two sub-components of early P3a, related to attention cognition and later P3b, related to memory storage. P3a is commonly observed in the frontal region of the brain where P3b is found to be maximum at temporal-parietal region [15], [16]. However, both peaks found in this study had their lowest amplitudes in the frontal area, in contradiction to the concept of P300 sub-component. Therefore, this suggests that there might be two or more potential types of advertising image stimuli that caused two different time of cognitive mechanism.

B. Clustering

Both K-means clustering and hierarchical clustering are algorithms that groups similar samples of similarity into groups based on the distance metrics. Therefore, we obtained the similar results from both methods. From the appearance of the peaks, two clusters might have been the best decomposition. Ideally, we expected to find that two clusters produced two clearly decomposable early and late P300 peaks. Unfortunately, the two sub-peaks of P300 were observed in one of the clusters when we clustered the images into two groups. Unexpectedly, we found two sub-peaks of P300 were observed in one of the clusters when we clustered the images into two groups. This implies that there were still a mixture of the images that caused the two P300 sub-peaks, and the two clusters were not successfully decompose two sub-peaks of

TABLE I
P-VALUES OF THREE P300 FEATURES OF TARGET AND NON-TARGET IMAGE PERCEPTION EXTRACTED WITHIN THE DURATION OF 0.30 TO 0.70 SEC OBTAINED FROM DEPENDENT T-TEST, WHERE *, **, AND *** DENOTE $p < 0.05$, $p < 0.005$, AND $p < 0.001$, RESPECTIVELY

Electrode conditions	Average amplitude	Maximum amplitude	Latency of maximum amplitude	Latency difference of target and non-target (msec)
Average across electrodes	2.221×10^{-10}***	1.072×10^{-10}***	0.003**	-0.050
F3	3.840×10^{-9}***	3.334×10^{-7}***	0.024*	-0.041
F4	3.697×10^{-8}***	1.199×10^{-6}***	0.021*	-0.042
C3	4.228×10^{-10}***	1.306×10^{-9}***	0.001**	-0.056
C4	5.179×10^{-9}***	3.002×10^{-9}***	0.001**	-0.055
Pz	3.267×10^{-8}***	6.143×10^{-9}***	0.002**	-0.051
PO7	2.487×10^{-10}***	3.857×10^{-11}***	0.451	-0.011
PO8	1.648×10^{-9}***	8.815×10^{-11}***	0.433	-0.014



(a) Elbow method (b) Silhouette method

Fig. 5. The results of k-value optimization in K-means clustering to determine the most suitable cluster number for P300 evoked during Coca-Cola advertisement image cognition task.

P300. In addition, we also observed that four cluster yield highest Silhouette score in silhouette analysis. However, one of four clusters, again still contained two P300 sub-peaks at the similar level of amplitude when we plotted the ERP of the images in that group. This also suggests that there were some typical types of target image that blended in that cluster. Therefore, the best possible decomposition of P300 peaks were considered to be three clusters of target image.

Among three clusters, one image cluster (group 2 of K-means and group 0 of hierarchical clustering results) evoked larger amplitude of early P300 which can be seen clearly in topography too. This suggests that the image in this cluster has better cognition with faster mechanism than the rest of the clusters. In our study, most of the images were in red, which is the symbolic colour of Coca-Cola with the name of the brand were clearly seen. Similarly, group 1 in in both clustering methods were evoked with a delay in latency as shown in topography in Fig. 7 and 8 but similar level of amplitude of the early P300. When we looked deeper into the images in this cluster, we noticed that both the images in this cluster were the same in both methods. They did not contain symbolic Coca-Cola ads colour and mostly filled with either black or white colour. However, we could still observe a mixture of P300 peaks in group 0 and group 2 obtained from K-means and hierarchical clustering methods, respectively. There might be some types of image hidden in the group that caused the sub-peaks or it could be noise. When we look into the images in this cluster, we could also see that there is a mixture of both symbolic colour of the brand and various colours of images and the majority of the images were the Coca-cola bottle, when compared to group 2 in K-means and group 0 in hierarchical clustering. It was also reported that the logo that participants disliked evoked lower N200 amplitude [17], the similar outcome might occur with the P300 amplitude as well, which means the P300 amplitude might depend on the preference of participants whether they like or dislike that advertising image. However, we still cannot conclude our above statements without further investigation.

C. Limitation and Future Work

In this study, we utilized the stimulus with different sizes and types (both only logo and poster with and without words) in our experiment. These factors could lead to more complexity in an image preference determination in cognition process which

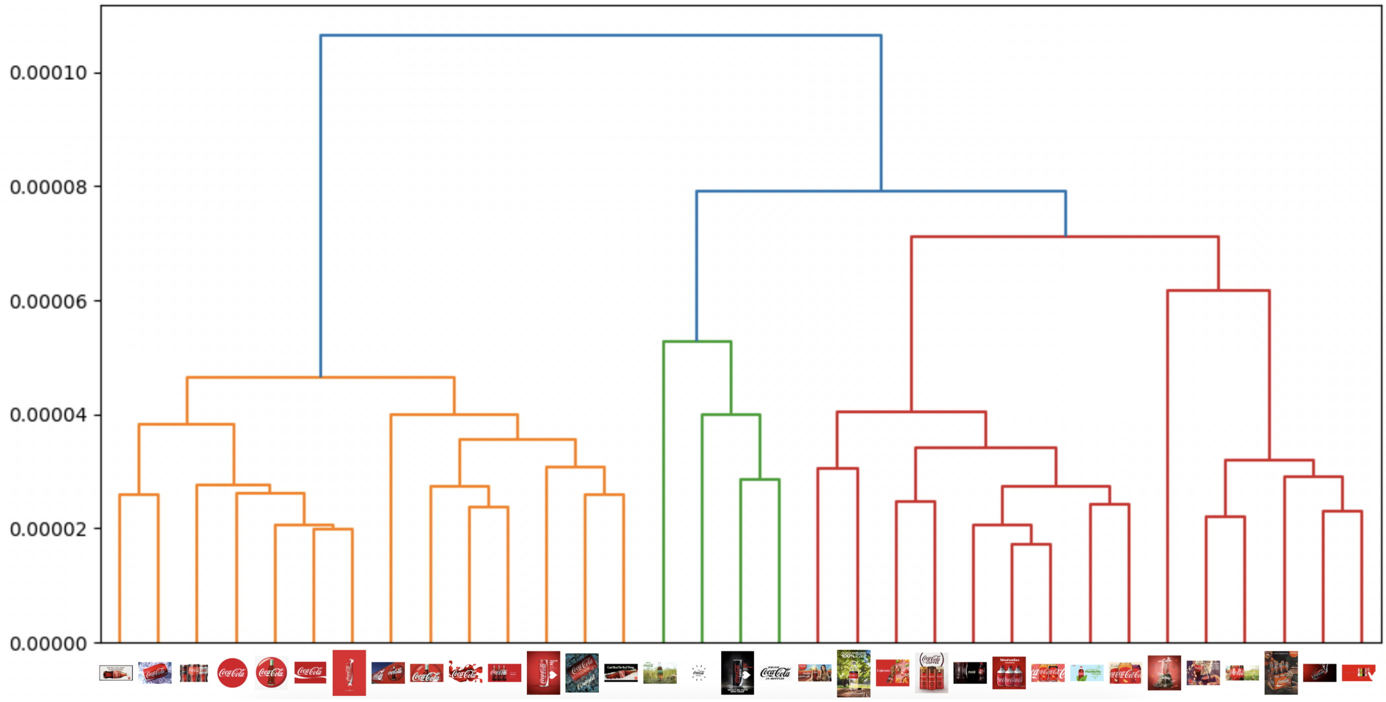


Fig. 6. The results of k-value optimization usingincrement algorithm in hierarchicalclustering todeterminethe most image cognition task

made the clustering becomes more difficult. This concern could be solved either designing the limitation and criteria for image stimulus selection or implement more complicated clustering algorithm. Moreover, we also have limitation in our experiment environment that our stimulus display was small and allowed the participants to see the environment of the experiment room behind the screen. This might distract them during the experiment. We had asked them to focus on the screen when performing the tasks, however, we have no prove that their eye gazes were on the screen all the time. To solve our mentioned limitations, we plan to extract the image features using image processing, clustering the images based on the image features and compared the results with the clusters based on EEG to identify the source of two P300 sub-peaks, and how P300 amplitudes change among each source. In addition, we plan to record eye tracking during the experiment too. With the help of eye tracking, we could confirm on how much the participants have been focusing on the screen and we could also gain the information on which part of the advertisement image draws their eyes at the first glance, and where their most interest region is.

Acknowledgment

This work was supported in part by JSPS Kakenhi 20H00235.

References

[1] N. Zmuda. (2009) Tropicana line's sales plunge 20% post-rebranding. [Online]. Available: https://adage.com/article/news/tropicana-line-s-sales-plunge-20-post-rebranding/135735
[2] C. Kirk and K. Berger, "Tropicana: Social media teach marketers a branding lesson," *Journal of Critical Incidents*, vol. 4, pp. 108–111, 01 2011.
[3] M. Bryant, "A study on rebranding strategies," 2013.
[4] G. S. Day, "The threats to marketing research," *Journal of Marketing Research*, vol. 12, no. 4, pp. 462–467, 1975.
[5] V. Fortunato, J. Giraldi, and J. Oliveira, "A review of studies on neuromarketing: Practical results, techniques, contributions and limitations," *Journal of Management Research*, vol. 6, p. 201, 04 2014.
[6] S. Kumar, "Neuromarketing: The new science of advertising," *Universal Journal of Management*, vol. 3, pp. 524–531, 12 2015.
[7] D. Ariely and G. Berns, "Neuromarketing: The hope and hype of neuroimaging in business," *Nature reviews. Neuroscience*, vol. 11, pp. 284–92, 03 2010.
[8] A. Bazzani, S. Ravaioli, L. Trieste, U. Faraguna, and G. Turchetti, "Is EEG suitable for marketing research? a systematic review," *Frontiers in Neuroscience*, vol. 14, p. 594566, 2020.
[9] M.-H. Lin, S. Cross, W. Jones, and T. Childers, "Applying eeg in consumer neuroscience," *European Journal of Marketing*, vol. 52, 01 2018.
[10] S. Lees, N. Dayan, H. Cecotti, P. Mccullagh, L. Maguire, F. Lotte, and D. Coyle, "A review of rapid serial visual presentation-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 15, 11 2017.
[11] G. F. Woodman, "A brief introduction to the use of event-related potentials in studies of perception and attention," *Attention, Perception, & Psychophysics*, vol. 72, no. 8, pp. 2031–2046, 2010.
[12] F. Guo, X. Zhang, Y. Ding, and X. Wang, "Recommendation influence: Differential neural responses of consumers during shopping online," *Journal of Neuroscience, Psychology, and Economics*, vol. 9, 02 2016.
[13] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.
[14] J. W. Y. Kam, Z. C. Irving, C. Mills, S. Patel, A. Gopnik, and R. T. Knight, "Distinct electrophysiological signatures of task-unrelated and dynamic thoughts," *Proceedings of the National Academy of Sciences*, vol. 118, no. 4, p. e2011796118, 2021.
[15] J. Polich and J. R. Criado, "Neuropsychology and neuropharmacology of p3a and p3b," *International Journal of Psychophysiology*, vol. 60, no. 2, pp. 172–185, 2006.
[16] J. Polich, "Updating p300: an integrative theory of p3a and p3b," *Clinical neurophysiology*, vol. 118, no. 10, pp. 2128–2148, 2007.

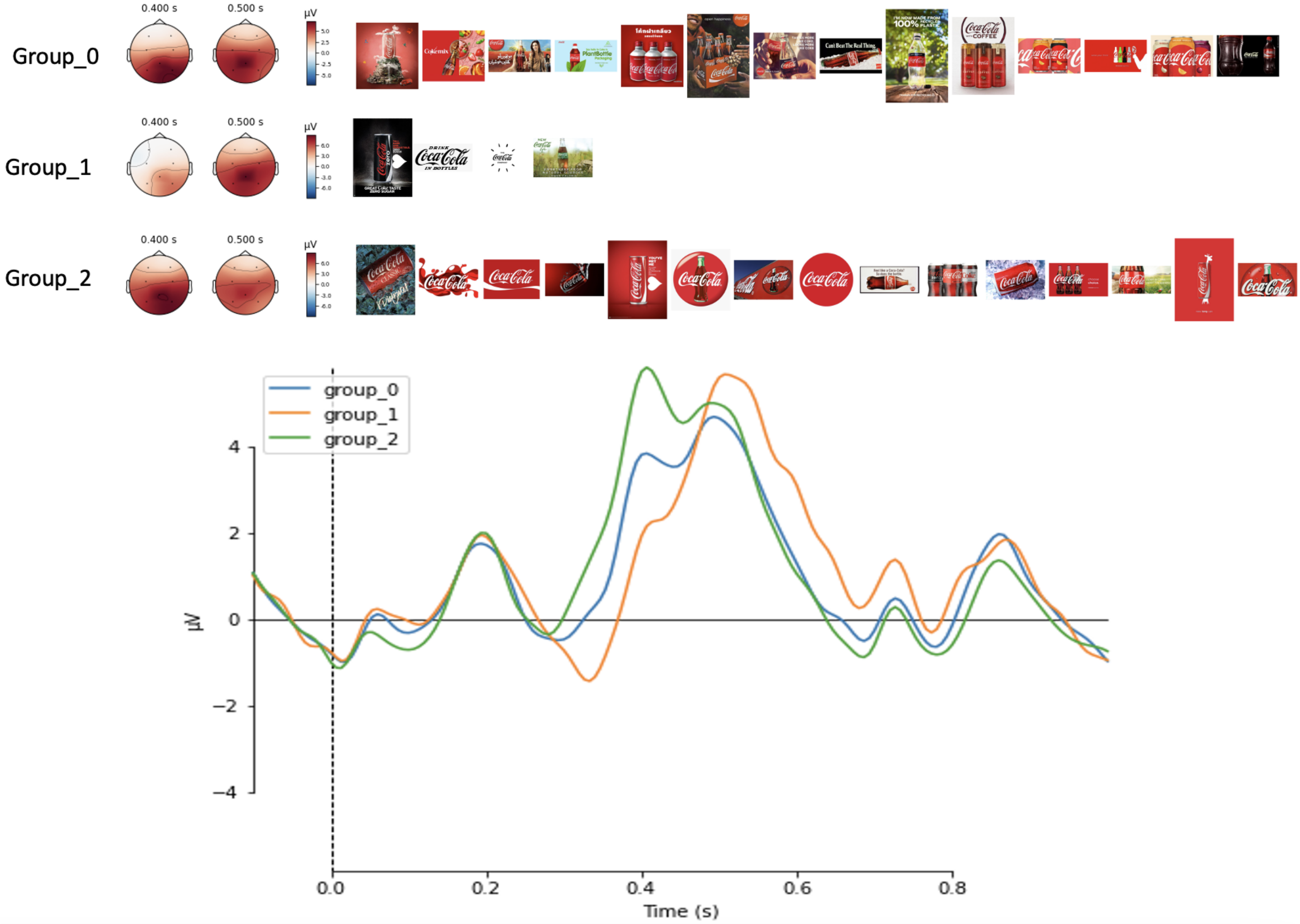


Fig. 7. The grand average ERP and topomap of all participants among Coca-Cola advertisement image groups when clustered by K-means clustering and the ads contained in each cluster

[17] T. C. Handy, D. Smilek, L. Geiger, C. Liu, and J. W. Schooler, "Erp evidence for rapid hedonic evaluation of logos," *Journal of Cognitive Neuroscience*, vol. 22, no. 1, pp. 124–138, 2010.

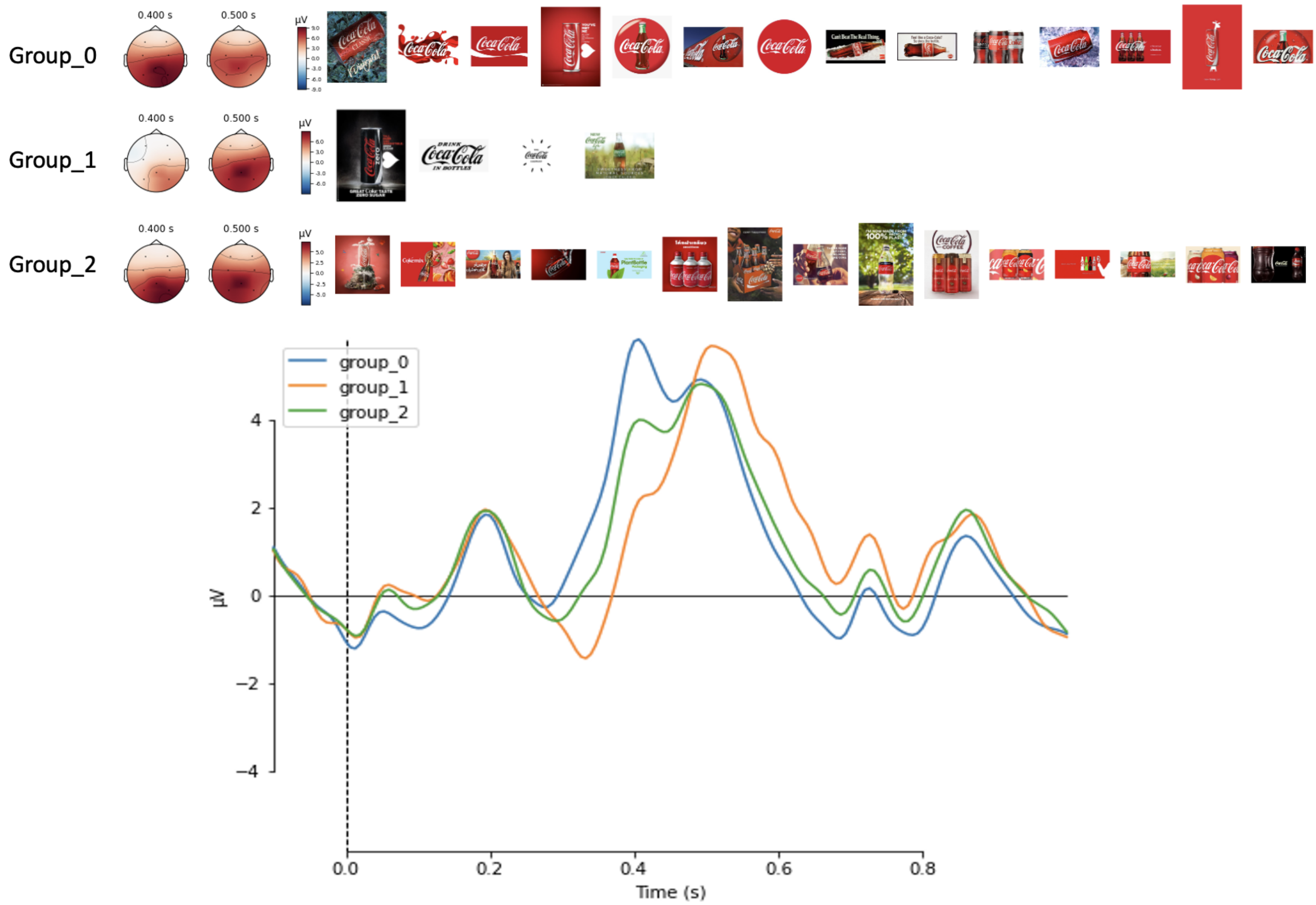


Fig. 8. The grand average ERP and topomap of all participants among Coca-Cola advertisement image groups when clustered by hierarchical clustering and the ads contained in each cluster