EXPLORING THE EYE TRACKING DATA OF HUMAN BEHAVIOUR ON CONSUMER MERCHANDISE PRODUCT

N.H. Baharom[1], S.R. Aid[1], M.K.M. Amin[1], S. Wibirama[2] and O. Mikami[3]

[1]Malaysian-Japan International Institute of Technology,
Universiti Teknologi Malaysia, 54100,
Kuala Lumpur, Malaysia.

[2]Department of Electrical Engineering and Information Technology,
Faculty of Engineering, Universitas Gadjah Mada,
Yogyakarta 55281, Indonesia.

[3]Department of Optical and Imaging Science & Technology,
School of Engineering, Tokai University,
Hiratsuka, Kanagawa 259-1292, Japan.

Corresponding Author's Email: [1]nurhellwa.baharom@graduate.utm.my

Article History: Received 27 December 2018; Revised 26 April 2019;
Accepted 13 October 2019

ABSTRACT: This article presents an exploration of the human eye tracking data towards consumer products. The study aim to investigate the data attributes of the cognitive processes and focused on the visual attention of the participants when choosing a shampoo brand which is commonly available in Malaysia. However, eye tracking datasets has a wealth of data on the eyes visual attention, fixation, saccade and scan path gaze. Therefore, this paper aims to solve this problem to minimize the datasets by using clustering machine learning approach. This is to observe the relation of these data attributes and possibly predict the possible solution contributing to cognitive processing. Tobii TX300 Eye-tracker was used in this experiment and the eyes tracking data were gathered particularly related to the eyes fixation and saccades by using the Tobii I-VT filter. Sixty subjects participated in this study. K-means clustering was used as statistical analysis to cluster the huge datasets from the eye tracking data. The relationship of the consumer cognitive processes with visual attention was understood when most of the participants chose the most popular shampoo brand such as Head & Shoulder. Further visual analysis on the data attributes results showed that K-means clustering has the potential to cluster and minimize the huge datasets and predicts consumer preferences.

KEYWORDS: *Eye Tracking; Exploring Data; Human Behavior; Visual Attention*

1.0 INTRODUCTION

The advancement of eye tracking technology has attracted many researchers to combine its usage with other equipment like Electroencephalogram (EEG) to understand the consumer behavior [1]. It is a potential strategy with this bio-equipment's when compared to traditional marketing. Both eye tracking and EEG have a huge data attributes since recording in any experiments will measures different parameters in milliseconds. This study therefore, focused on huge eye tracking data attributes and clustered the data as an initial step to investigate and determine the relationship of the visual attention data attributes and consumer preferences.

Big data was the information assets [2] and technologies that require new forms of integration [3] to disclose hidden value of complex datasets into value. In order to develop some more complex statistical methods to analyze those massive data, it requires a deep understanding. For example, infants retain memories of visual experience over time where such early experience can have cascading effects on human development [4]. The final outcome will be change due to different decisions, leading to different results and accordingly different explanations [5]. With such discovery, eye movements are proven to be useful in accessing cognitive process in young infants.

In recent years, there has been an increasing amount of literature on eye tracking study on purchasing behavior in different method such as [6] using a conditional logit analysis method to examine consumers' visual attention. In a study which set out to determine consumer's behavior [7-8], found that eye tracking data could predicts consumer's choice by using ANOVA analysis. Some analysts such as [9] used Generalized Linear Mixed Models (GLMM) and U-shape curve function in [10] to investigate the factors of affecting the consumers' visual attention when purchasing the product. Previous studies have shown some of the limitations such a method cannot be understood thoroughly on how to manage a wealth datasets on predicting consumer preferences. Based on the reason, the objectives of this study are to observe and understand the relationship between the data attributes and visual consumer preferences.

2.0 METHODOLOGY

2.1 Participants and Equipment

Sixty students of Malaysia-Japan Institute of Technology (MJIIT) were recruited for this experiment. All participants were aged between 20 and 25 years old with number of 30 males and 30 females. They were all mentally and physically healthy and without any diseases that related to eye. The Tobii TX-300 eye tracker was used as a main method and the experiment was conducted inside the lab with the temperature of 25 degree Celsius. During experiment, the calibration was needed for every participant to ensure that the data collected is with least noise and best quality. On average, each participant took 5 minutes to complete the experiment.

2.2 Data Analysis

The data obtained from this experiment are mainly analyzed using the visual data analysis generated from the Tobii Eye-tracker. R language was used for further statistical analysis. This language is an open source programming and software environment and offers a lot of machine learning. It has detailed advanced tasks for analyzing big data, which supports high performance statistical and graphical computation [11]. In this experiment, the NbClust of the R packages was used in order to determine the relevant number of clusters [12]. Hence, as it determined the optimal number of clusters in a data set, it can offers the best clustering pattern from different results to the user and scan path [13].

3.0 RESULTS AND DISCUSSION

3.1 Results of Clustering Analysis on Eye Movement Data

The number of cluster, K needs to be defined before the k-means algorithm applied for clustering. Nbclust package in R was used to determine the optimal clusters. K-means algorithm was well known for its efficiency in clustering huge data sets due to its low computational cost and easily parallelized process. However, various extensions of k-means algorithms have been developed to speed up the clustering process. For instance, there were kernel k-means [14], spherical k-means [15], and fuzzy c-means [16]. Clustering techniques have been used for solving several important problems [17]. The range of the number of clusters was set in between 2 and 15 among 30 indices. Among all indices, 5 proposed 4 clusters was the best number of clusters. Therefore, the best number of clusters selected was 4. The result of determine the optimal cluster was shown in the Figure 1.

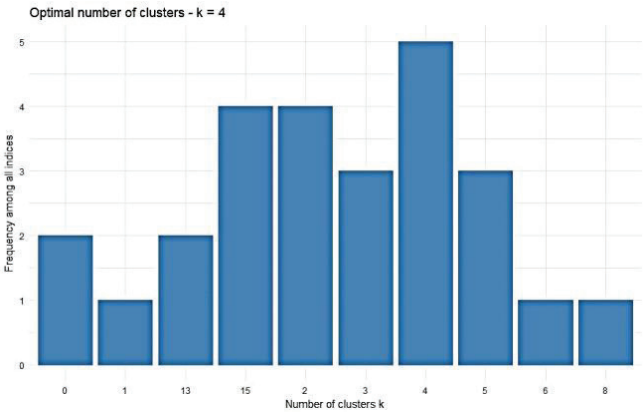

Figure 1: Determining optimal cluster using Nbclust

In Figure 2, 4 clusters were determined in the result of the K-means clustering by using Nbclust package. The colors represent the different clusters and the cross indicates the centre of each cluster. The behavior of each cluster is determined by the scan paths that nearest to the centre of each cluster. Figure 2 also showed 25 represented the blue cluster, 6 represented the red cluster, 38 represented the green cluster and 31 represented the black cluster. The scan path patterns for every participant were determined and divided into four groups.

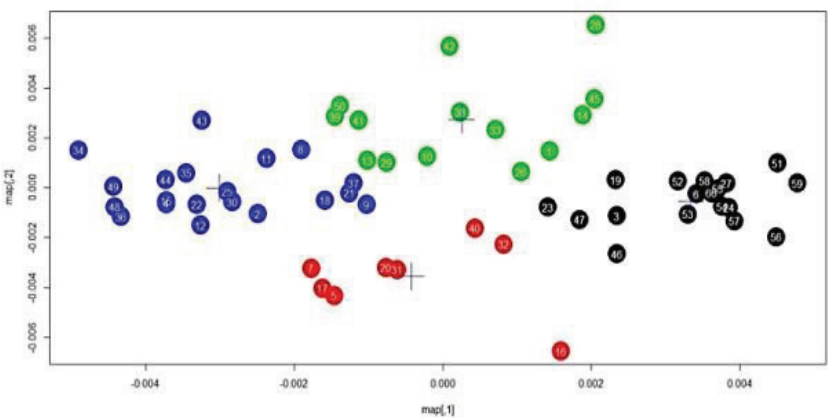

Figure 2: Clustering result using k-means algorithm with cluster number of 4

Most of the participants were grouped in blue cluster (20 participants), followed by black cluster (18 participants), green cluster (14 participants) and red cluster (8 participants). Figure 3 shows the gaze pattern of participant N25 while Figure 4 shows the simplified scan path of the participant N25. The scan path pattern is measured in clockwise scanning after looking at the center of the image. In this

cluster, most of the participants choose B (Head & Shoulders) (16 participants), followed by A (Pantene) (3 participants) and F (Johnson's Baby) (1 participant).

Figure 3: Gaze pattern of participant N25

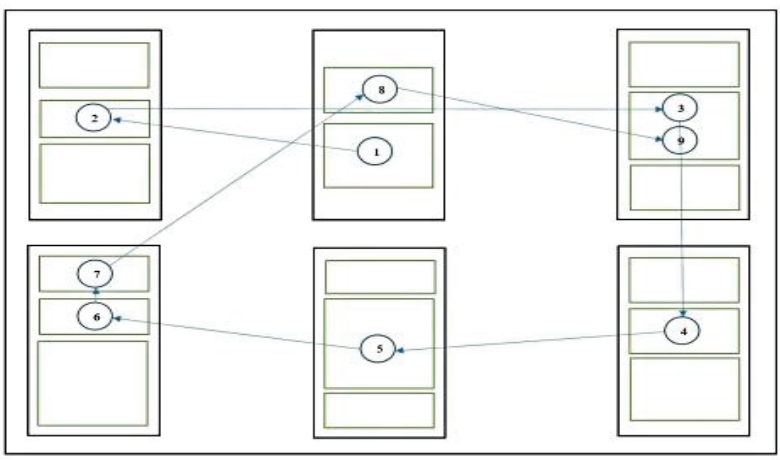

Figure 4: Simplified scan path of the participant N25

Figure 5 shows the gaze pattern of participant N6 while Figure 6 shows the simplified scan path of the participant N6. The scan path pattern is in Z scanning after looking at the center of the image. In this cluster, most of the participants choose B (Head & Shoulders) (13 participants) followed by A (Pantene) (3 participants) and F (Johnson's Baby) (2 participants).

Figure 5: Gaze pattern of participant N6

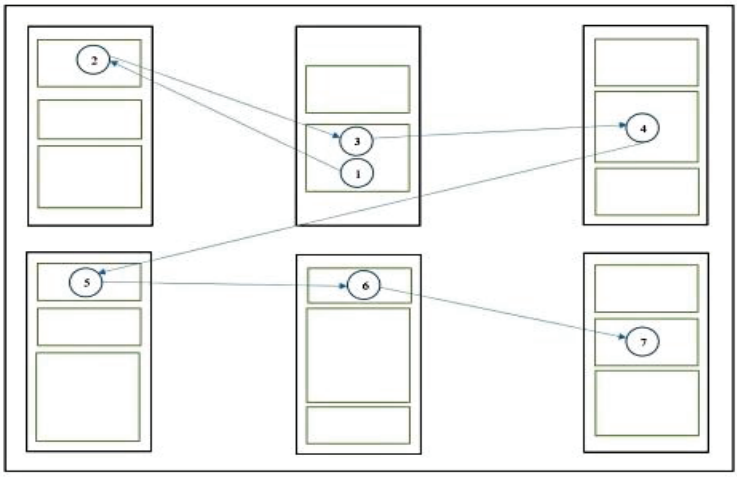

Figure 6: Simplified scan path of the participant N6

Figure 7 shows the gaze pattern of participant N38 while Figure 8 shows the simplified scan path of the participant N38. The scan path pattern scanned all the products twice while the fixation duration is short. In this cluster, most of the participants choose B (Head & Shoulders) (7 participants) followed by A (Pantene) (6 participants) and F (Johnson's Baby) (1 participant).

Figure 7: Gaze pattern of participant N38

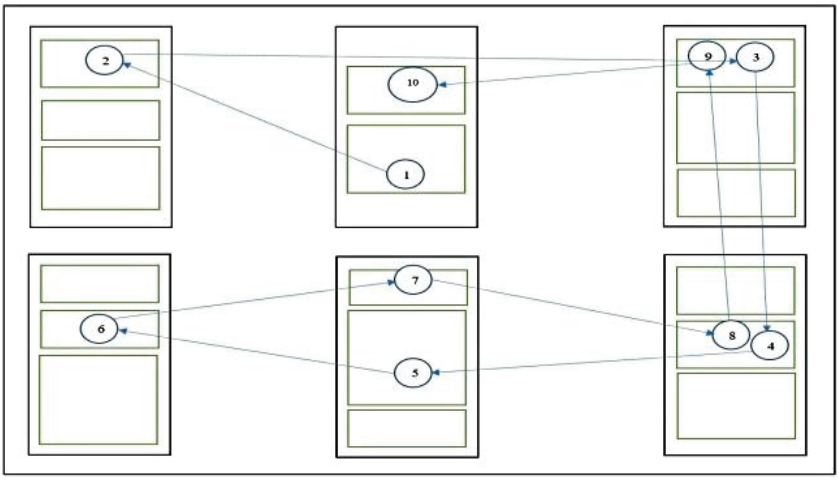

Figure 8: Simplified scan path of the participant N38

Figure 9 shows the gaze pattern of participant N31 while Figure 10 shows the simplified scan path of the participant N31. The scan path pattern scanned a certain product only and the fixation duration is long. In this cluster, most of the participants choose A (Pantene) (3 participants) followed by B (Head & Shoulders) (2 participants), E (Clear) (2 participants) and F (Johnson's Baby) (1 participant).

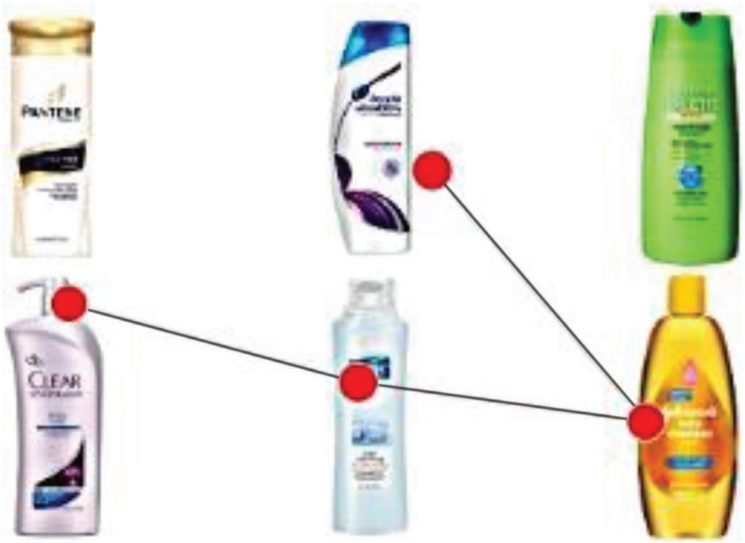

Figure 9: Gaze pattern of participant N31

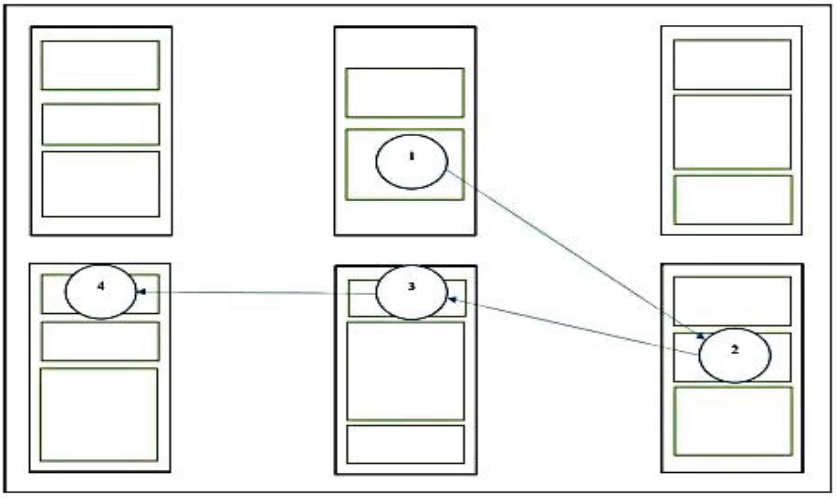

Figure 10: Simplified scan path of the participant N31

3.2 Critical Discussion

Figure 3 until Figure 10 concluded that Head & Shoulders was the famous and common shampoo brand chosen by young age of participants followed by Pantene and Johnson's Baby. It also concluded that branding was the most consumer preferences to be considered for purchasing. These findings also supported the study [18] that brands were vital factor in guiding the decisions during purchase process and

[19] also stated that brand was the factor which has and impact on the consumer preferences. This study confirmed that the consumer preferences can be forecasted through eye tracking datasets and also supporting the study in [20] that visual attention is the most related parameter with the cognition of the products. In future, this study will be continued by increasing the participant to be at least, 100 participants with different background and the eye tracker glasses will be applied outside the laboratory.

4.0 CONCLUSION

The main goal of the current study was to determine the relationship of data attributes with the visual consumer preferences. This study has found that generally K-means clustering technique could be useful to minimize the huge datasets of eye tracking (as a visual equipment) in order to classify the consumer preferences (branding). The present study confirms that the objectives were achieved as branding was the consumer preferences. In the industrial contribution, eye tracking may be useful as a sturdy benefit as eye movement is the first being faster than any other input media. Hence, marketing success of businesses can be continuously enhancing such as the products packaging with brand awareness strategies as well as brand loyalty improvement.

ACKNOWLEDGMENTS

The author would like to thank members of Bio cognition Laboratory of Bio-inspired System Technology research group, Malaysia-Japan International Institute of Technology (MJIIT). This work was supported under UTM research grant PY/2017/01721.

REFERENCES

[1] R. N. Khusaba, C. Wise, S. Kodagoda, J. Louviere, B. Kahn and C. Townsend, "Consumer neuroscience: Assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking", *Expert Systems With Applications*, vol. 40, no. 9, pp. 3803-3812, 2013.

[2] M. Raschke, T. Blascheck and M. Burch, "Visual Analysis of Eye Tracking Data," in *Handbook of Human Centric Visualization*, W. Huang, Ed. Switzerland AG: Springer, 2014, pp. 391-410.

[3] A. De Mauro, M. Greco, and M. Grimaldi, "What is big data? A consensual definition and a review of key research topics," AIP Conference Proceedings, vol. 1644, no. 1, pp. 97-104, 2015.

[4] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. U. Khan, "The rise of 'big data' on cloud computing: Review and open research issues", *Information Systems*, vol. 47, pp. 98-115, 2015.

[5] R. L. Fantz, "Visual experience in infants: Decreased attention to familiar patterns relative to novel ones", *Science*, vol. 146, no. 3644, pp. 668-670, 1964.

[6] R. Pieters and L. Warlop, "Visual attention during brand choice: The impact of time pressure and task motivation", *International Journal of Research in Marketing*, vol. 16, no. 1, pp. 1-16, 1999.

[7] S. Goyal, K. P. Miyapuram and U. Lahiri, "Predicting Consumer's Behavior Using Eye Tracking Data," in Second International Conference on Soft Computing and Machine Intelligence, Hong Kong, 2015, pp. 126-129.

[8] Q. Wang, M. Wedel, L. Huang and X. Liu, "Effects of model eye gaze direction on consumer visual processing: Evidence from China and America", *Information and Management*, vol. 55, no. 5, pp. 588-597, 2018.

[9] K. Gidlof, A. Anikin, M. Lingonblad, and A. Wallin, "Looking is buying. How visual attention and choice are affected by consumer preferences and properties of the supermarket shelf", *Appetite*, vol. 116, pp. 29-38, 2017.

[10] R. V. Menon, V. Sirgurdsson, N. M. Larsen, A. Fagerstrom, and G. R. Foxall , "Consumer Attention to price in social commerce: Eye Tracking Patterns in Retail Clothing", *Journal of Business Research*, vol. 69, no. 11, pp. 5008-5013, 2016.

[11] M. Charrad, N. Ghazzali, V. Boiteau, and A. Niknafs, "NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set", *Journal of Statistical Software*, vol. 6, no. 61, pp. 1-36, 2014.

[12] T. Von der Malsburg and S. Vasishth, "What is the scanpath signature of syntactic reanalysis?", *Journal of Memory and Language*, vol. 65, no. 2, pp. 109-127, 2011.

[13] I. S. Dhillon, Y. Guan and B. Kulis, "Kernel k-means: spectral clustering and normalized cuts", in Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, USA, 2004, pp. 551-556.

[14] I. S. Dhillon and D. S. Modha, "Concept Decompositions for Large Sparse Text Data using Clustering", *Machine Learning*, vol. 42, no. 1-2, pp. 143-175, 2001.

[15] J. Bezdek, *Pattern recognition with fuzzy objective function algorithms.* New York: Plenum Press, 1981.

[16] A. McCallum, K. Nigam and L. H. Ungar, "Efficient clustering of high-dimensional data sets with application to reference matching", in the Proceedings of the sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Boston, USA, 2000, pp. 169-178, 2000.

[17] T. S. Madhulatha, "An Overview on Clustering Methods", *IOSR Journal of Engineering*, vol. 2, no. 4, pp. 719-725, 2012.

[18] A. Mishra and O. Dutt, "Importance of Brand Name in Consumer Decision Making Process", *International Research Journal of Management Sociology and Humanity*, vol.5, no. 2, pp. 873-890, 2014.

[19] H. H. Chovanova, A. I. Korshunov and D. Babcanova, "Impact of Brand on Consumer Behavior", *Procedia Economics and Finance*, vol. 34, pp. 615-621, 2015.

[20] H. Zamani, A. Abas and M. K. M. Amin, "Eye Tracking Application on Emotion Analysis for Marketing Strategy", *Journal of Telecommunication, Electronics and Computer Engineering*, vol. 8, no.11, pp. 87-91, 2018.