# Practical approach to Eye Tracking Saliency on Commercial Advertising Video using Clustering

Hyunjong Lee, Jeehyun Bae, Jooyoung Oh, Kyuho Son, Youngjin An

***Abstract*—Commercial video market is the one of the most fast growing market these days. To evaluate the efficiency of a commercial video, analysis on response of viewers is essential. Gaze movement is the most representative factor among the responses. Extracting frequent gaze movement patterns in the responses and providing them as a feedback will produce lots of new opportunities in the market. We implemented clustering that classifies the gaze movement data into different patterns with K-Means algorithm, and also developed visualization tool for the clusters. We found that K-Means algorithm provide a meaningful features on clustering gaze movement data, but there also was a limitation of cluster instability due to tiny size of data and random initialization of clusters in K-Mean algorithm.**

***Keywords*—*Gaze, Eye tracking, Clustering, K-Means, Saliency, Commercial Video.***

## I. Introduction

Developing a tool for evaluating and analyzing the process of a commercial advertisement can provide the feedback for better production and efficiency in both time and budget. There has been several approaches for analyzing the users behavior while watching advertisement to find out an optimal experience and maximum effect on viewer. Johann Schrammel introduced multi-sensor eye tracking design for understanding attention of the user on outdoor advertisment [1]. For the evaluation, Wedel have assessed marketing efficacy of the visual attention tracking [2]. However, current research has its limitation in gathering the direct feed back from the audience with mass amount of data. The self-report or field research in eye tracking viewer's attention has its limitation in admitting big data. Since company and advertisement production cannot afford the testing real-time feedback of the viewer in a large scale, it is crucial to simulate the viewers approach when only limited feedback is allowed from the directors and pre-post production company.

### A. The Object-level video advertisement

Current studies can be identified into 4 types of video advertisement(mostly on online) as a framework. Most current studies are 1) locating ads in appropriate place, 2) Text based advertisement, 3) Video segment level advertising, 4) object-level video advertisement. Locating advertisement utilizes the blank space of the video advertisement and attach the brand logo or image. Text based advertisement matches related words searched by the user and presents the advertisement. thirdly, segment level advertisement analyzes the current video and allow the most related image into the scene. The object-level video advertising (OLVA) which regards object as a main component of the video advertisement analyzation [3]. In this study our asumption was mainly focused on the moving object in the video; which gains the most part of the viewer's gaze. Then regarded it as a main target for the whole advertisement. The data colledcted for the clustering was used to represent the one moving object (with and without the sound).

### B. Eye Tracking Saliency

Audiences attention can be estimated and analyzed by tracking the Eye Tracking Saliency. In this paper, we present an application which use the machine learning algorithm to compute the mass data for the analyzation of the gaze sample data. The system utilize the Aggregaze [4] which estimate an attention of audience on public display by single off-the-shelf camera attached to the display. Machine Learning is used for augmenting the users current gaze without any personal calibration. Aggregaze allows uncalibrated, inaccurate gaze estimates of multiple users to joint attention estimates. Using the Aggregaze data, our project deals with the visualization and analytic on the commercial advertising video. In this paper, our aim is to estimate the viewers attention and patternize the the viewer's gazing traits. The whole process is to find out whether the producer's goal is achieved, which is to deliver the right message that advertisement intended via audience's gaze. This practical approach allows to analyze viewers' gaze patterns in respect of time. The obtained data can be used to evaluate different types of video production.

## II. Experimental Setup

Total process of experiment consist of clustering, verification of clustering and visualization of clusters in user interface. We can get cluster indices of each samples and the mean gaze record of each clusters in clustering stage. Then, user interface for visualization will illustrate the mean records of the clusters with them on the video material played. Clustering stage Verification stage is not included in user's process, but was used to verify how the clustering performs well.

### A. Clustering

Clustering was performed in Python, Windows. We used PyCharm, Anaconda with SciPy and Scikit learn library.

We chose K-Means as the clustering algorithm. K-Means is an algorithm that clusters data to K classes with Euclidian distance. [5] Data of gaze point record is 2-D location based. So, in analyzing the data, similarity of Euclidian distance can

Source code available at
https://github.com/AblemanAS/cs492_2017spring_gazeclustering

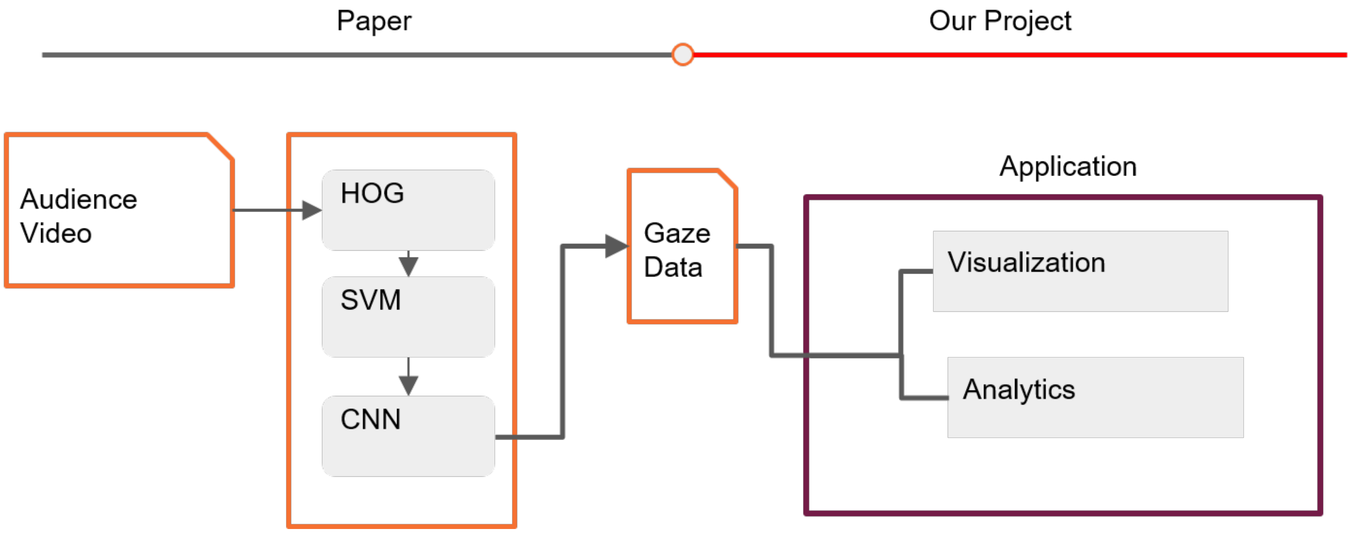



Fig. 1: Project Scheme

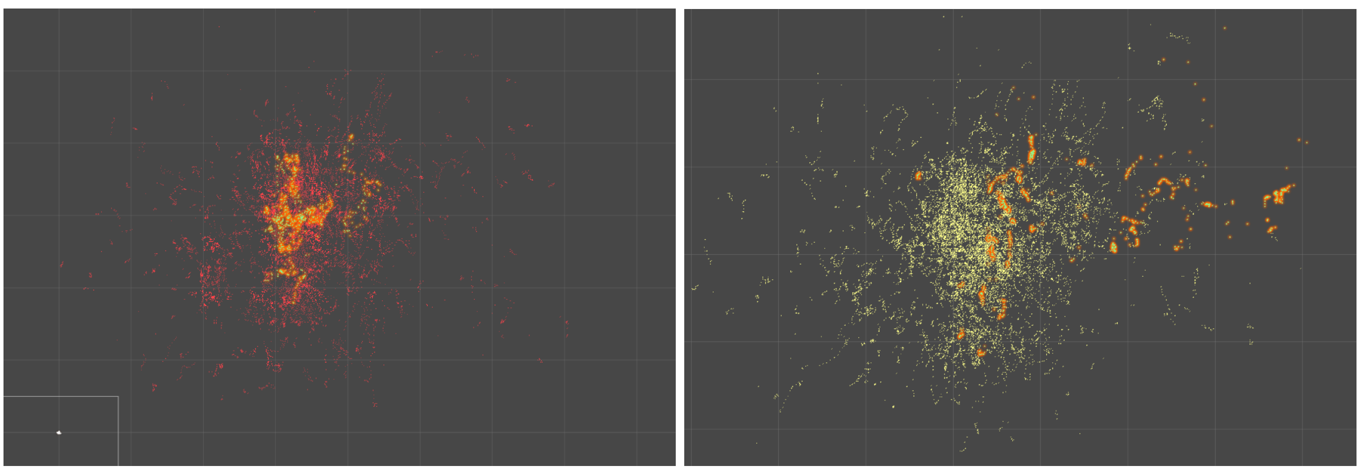


Fig. 2: Clustering Verification

be a good factor in clustering. This is the reason why we chose K-Means as the clustering algorithm.

To select the optimal number of K, we calculated silhouette score of each K every iteration. Silhouette score is a succinct graphical representation of how well each object lies within its cluster. [6]

### B. Verification of Clustering

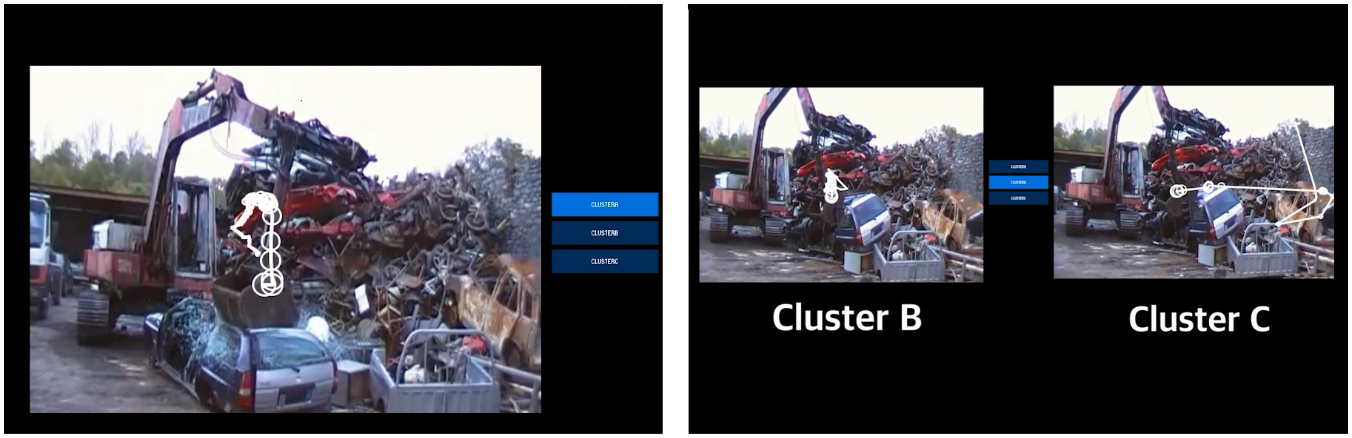
Unity(5.3.0) was used for verifying clustering results. Time stamps of each data set are instantiated as sphere objects and located after scaling by x, y value. We hierarchically construct these objects by data set or cluster for changing visibility easily.

We can judge clustering results after scattering objects, choosing cluster or data sets which are included cluster and changing visibility by enabling or disabling the hierarchy. We also proved results by sequence of time stamps because they are instantiated by time flow.

Figure 2. shows scattered gaze points of each clusters. Highlighted points are the records of the mean of clusters. The other points are the records of samples those are classified into the clusters.

### C. User Interface and Visualization

User interface was developed with processing. [7] The interface has the feature that reads the clustered data and visualize them with timed gaze movement path. User interface is introduced with 3 maximum cluster choices. Each of the clusters are arranged in priority order. As shown in the Figure 3. (a), the data shows the most probable pattern of a viewer with the sample data. In Figure 3. (b), it shows different patterns of the gaze attention with lower probability. The video




(a) (b)

Fig. 3: (a) Cluster Mean Visualization (b) Different Patterns in Different Clusters

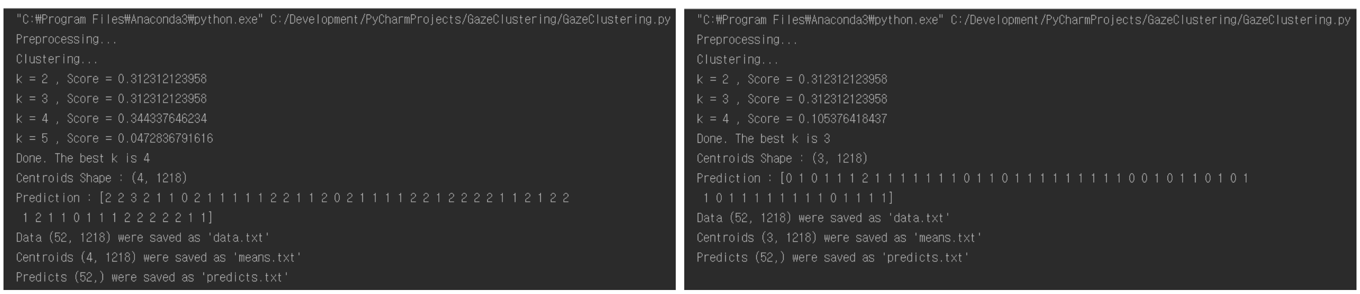



Fig. 4: Clustering Instability

clip and gaze pattern in synchronized in real-time to show the exact moment of the viewer's point. The gaze line is design to show the pattern in both time and direction. Firstly, as the point moves to the new direction the circle shape emerged on the video. The radius shape visualizes how long the viewer's gazing point stayed at a static position. Secondly, as the next point is predicted on the image location, white line is drawn to show the time it took from past gaze point to present point. The program provides comfort on watching different gaze patterns on a material video. Also, viewing gaze movements of every individual participants was not implemented due to point out the main function of the program.

## III. Result

For the data set, 3 to 4 clusters were extracted from about 50 samples. And 80% of samples gathered in 2 clusters. Final number of clusters K varied 3 to 5 by random initial points in K-Means Algorithm. Figure 4. shows the instability. In the visualization, the program showed different patterns by different clusters. Unlike to cluster A, cluster B was somewhat distracted. Cluster C tended to watch the video in a wide view.

## IV. Conclusion

Each clusters represent different patterns of gaze movement appositely. And for instability of K different initial points in K-Means algorithm, this can be perceived as weakness in catching out-liar instances when they are frequent. the larger data set would resolve the problem.

## V. Discussion

K varied with different initial points in K-Means algorithm. As we mentioned, it would become stable in a larger data. But not all data sample can be large enough. The more researches on stability on small data set are needed. Like the inspiration of this paper, combination with Aggregaze to apply on publicly displayed videos watching the same clip multiple times would provide full fledged story line. And if we can classify those clustered pattern into acknowledged classes, we would be able to extract user-class-adaptive thumbnail image of video and expose it to the users of the class.

### A. Expansion to Commercial Video

Returning to the start point, the terminal target of our project is commercial video analysis. Among them, the easiest targets are public displays such as the one in theater or on sky scrapers because of those comforts of appliances of gaze trackers. So these days there are many trials of analysis on commercial video using human saliency. [8] The correspondent part of the article is mainly on moving objects, but the author says that these analysis can be also applied on text-based messages and so on. Our research to be expanded to real commercial video, further studies on clustering of gazes on multiple moving objects is needed. When these are done, we would be able to get much closer to practical usage of our research on commercial video analysis.

## APPENDIX A
## GAZE MOVEMENT VISUALIZATION PROGRAM DEMONSTRATION

https://youtu.be/Xu1JhpPNdKI

## REFERENCES


[1] J. Schrammel, E. Mattheiss, S. Döbelt, L. Paletta, A. Almer, and M. Tscheligi, "Attentional behavior of users on the move towards pervasive advertising media," in *Pervasive Advertising*. Springer, 2011, pp. 287–307.

[2] D. McDuff, "New methods for measuring advertising efficacy," *Digital Advertising: Theory and Research*, p. 327, 2017.

[3] H. Zhang, X. Cao, J. K. Ho, and T. W. Chow, "Object-level video advertising: An optimization framework," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 520–531, 2017.

[4] Y. Sugano, X. Zhang, and A. Bulling, "Aggregaze: Collective estimation of audience attention on public displays," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 2016, pp. 821–831.

[5] S. Lloyd, "Least squares quantization in pcm," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[6] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.

[7] P. Foundation. Processing. [Online]. Available: https://processing.org/

[8] F. Alt, *A design space for pervasive advertising on public displays*. Lulu.com, 2013.