

Preliminary Analysis of Video Watching Interest Using Fixation Frequency and Pupil Size

CHANNA VAN^{†1}SUGANUMA MUTSUMI^{†2}WATARU KAMEYAMA^{†3}

We present the analysis of video watching interest using fixation frequency and pupil size. The experiments are conducted by asking subjects to watch 3 types of videos, including comedy, talk show and self-made video, selected from YouTube. By analyzing the distribution of the fixation frequency and the pupil size, we have found that the higher fixation frequency and the larger pupil size are observed in case of watching the video which is rated to be interesting by subjects. This result shows the possibility to utilize the fixation frequency and the pupil size for analyzing the interest level of subjects while watching videos.

1. Introduction

Eye movements are believed to be able to express human emotion while watching objects and scenes based on various features such as gaze, pupil size, fixation, blink and so on[1][2][3][4]. However, in video watching, it is difficult to define features to express human interest due to the nature of video which is a set of non-static images. It means that regions of interest (ROI) change from scene to scene in video. Therefore, extracting the pattern of gaze paths, gaze points and fixation positions is very difficult because gaze paths and gaze points and fixation positions vary to the position of the ROI of video. In the previous studies on eye movement and video watching, they mainly focus on the usage of ROI for studying the interest level such as in [5]. In [6], [7] and [8], they propose to use eye movement features such as pupil sizes, blinks and gaze points cooperated with other human bio-signals such as heart rates, facial movements and head movements in order to capture the interesting part of video for video summarization.

In this study, we explore the relationship of video watching interest and the ROI-independent features of eye movement to measure video watching interest level. And we propose the utilization of fixation frequency and pupil size. The fixation frequency is the number of the extracted fixations during certain watching duration. The fixation frequency and the pupil size are independent from the ROI that make them the best candidate features for analyzing video watching interest.

2. Experimental Methodology

Experiments are conducted by asking subject to watch 3 types of video in 3 watching sessions. Each session lasts 5 minutes. The eye movement and the pupil size are recorded using an eye tracker device.

2.1 Apparatus

To capture eye movement information, we have used an eye tracker device, Mirametrix S2[4]. The specification of the device is shown in Table 1. A 19 inches LCD monitor has been used for displaying videos.

Table 1 Mirametrix S2 Specification

Accuracy	< 1° of visual angle
Head Motion	25 x 11 x 30cm
Data Rate	60 Hz
Binocular Tracking	Yes
Tracking Type	Bright Pupil

2.2 Subjects

Five volunteered subjects, one female and four males, participated in this experiment. Their ages ranged from 23 to 28. All subjects were confirmed that they had never watched the selected videos before. Furthermore, subjects were not aware of the research purpose and the results of the experiment.

2.3 Videos

We have selected 3 different types of video: comedy, talk show and self-made video from Youtube with the length of 5 minutes. We expect that these 3 types of video provide different watching interest level. Our hypothesis is that the comedy should provide the highest watching interest, the talk show should provide the moderate watching interest, while the self-made video should provide the lowest watching interest.

2.4 Data Collection

The open source gaze analyzer Ogama[5] was used to record the data of eye movement information from the eye-tracking device into database. The data of the pupil sizes and the fixations were exported from the database using Ogama. The fixations were extracted at the sampling rate of 60Hz.

2.5 Experimental Procedure

Subjects were instructed to watch 3 types of video in 3 different sessions. Each session lasted 5 minutes.

Before each watching session, eye-tracking calibration was carried out to calibrate the eye behaviors of each subject. After each watching session, subjects were asked to provide watching interest level of each video based on the scale from 1 to 5 where 1 defines the lowest interest and 5 defines the strongest interest. While watching videos, the eye-tracking device tracked subjects' eye movement as well as pupil size, and the eye movement data were recorded in database using Ogama.

^{†1} GITS, Waseda University

^{†2} GITI, Waseda University

^{†3} School of Fundamental Sc. and Eng., Waseda University

3. Results and Discussion

3.1 Watching Interest Level Rating

Table 2 shows the watching interest level rating of all subjects. On average, the comedy video received the most favorable rate of 3.8, and the talk show received 3.2, while the self-made video received the lowest rate 2.

Table 2 Subjects' Interest Level Rating

Subject	Comedy	Talk Show	Self-made
Subject1	4	2	2
Subject2	4	2	1
Subject3	4	5	4
Subject4	4	3	1
Subject5	3	4	2

3.2 Fixation Frequency

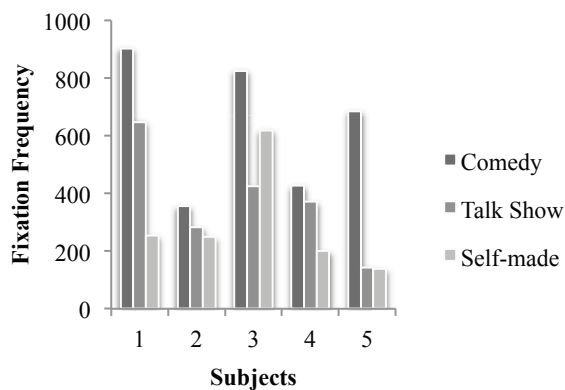


Figure 1 Fixation Frequencies of Each Subject for Each Video

The result of fixation frequencies in each 5 minutes video as shown in Figure 1 shows the significant differences between the highest rated video and the lowest rated video. In general, the high rated comedy video corresponds to the highest fixation frequencies for all subjects; the lowest rated self-made video has the lowest fixation frequencies except in the case of subject 3 where it increases due to his/her high rating.

3.3 Pupil Size

Similar to the fixation frequency results, the average pupil size results show that larger pupil size is observed in case of comedy comparing to other types of videos as shown in Figure 2. However, no significant differences of pupil size can be observed between the talk show and the self-made video, which are almost the same.

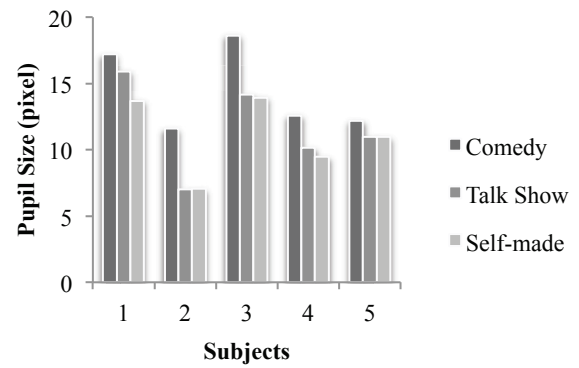


Figure 2 Average of Pupil Size of Each Subject for Each Video

3.4 Fixation Frequencies and Fixation Durations

Figure 3, 4, 5, 6 and 7 show the distributions of fixation frequencies per fixation duration of subject 1, 2, 3, 4 and 5, respectively. Based on these distribution graphs, we can observe that the high concentration of fixations appears between the ranges of 60ms to 200ms of the fixation duration where its maximum range is between 80ms to 100ms. The significant differences of the fixation frequencies are also observed in the maximum range for each subject.

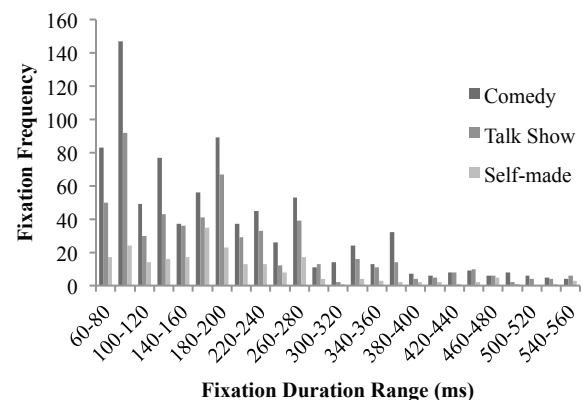


Figure 3 Distributions of Fixation Frequencies per Fixation Duration of Subject 1

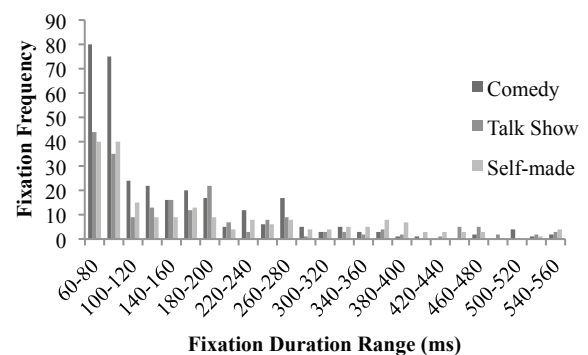


Figure 4 Distributions of Fixation Frequencies per Fixation Duration of Subject 2

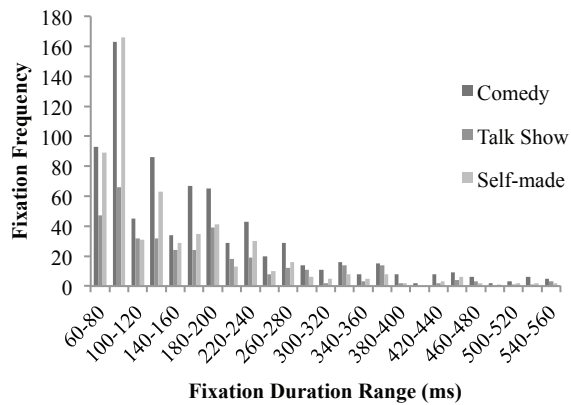


Figure 5 Distributions of Fixation Frequencies per Fixation Duration of Subject 3

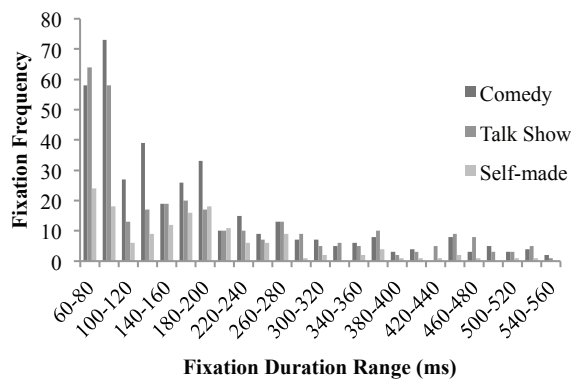


Figure 6 Distributions of Fixation Frequencies per Fixation Duration of Subject 4

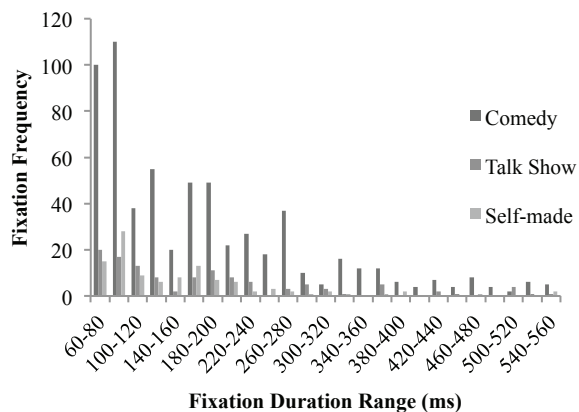


Figure 7 Distributions of Fixation Frequencies per Fixation Duration of Subject 5

3.5 Fixation Frequencies and Pupil Sizes

In this subsection, we present the relationship of the fixation frequencies and the pupil sizes. Figure 8, 9, 10, 11 and 12 show the distributions of the fixation frequencies per pupil size of subject 1, 2, 3, 4 and 5, respectively. Based on these figures, the high fixation frequencies and the large pupil size can be observed in case of the comedy in all subjects. This shows the

relationship of the fixation frequencies and the pupil sizes in the case of high rated video.

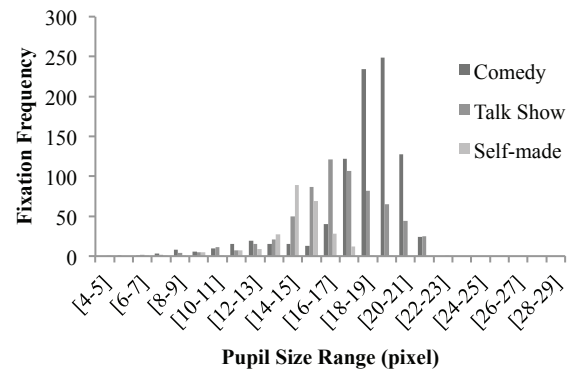


Figure 8 Distributions of Fixation Frequencies per Pupil Size of Subject 1

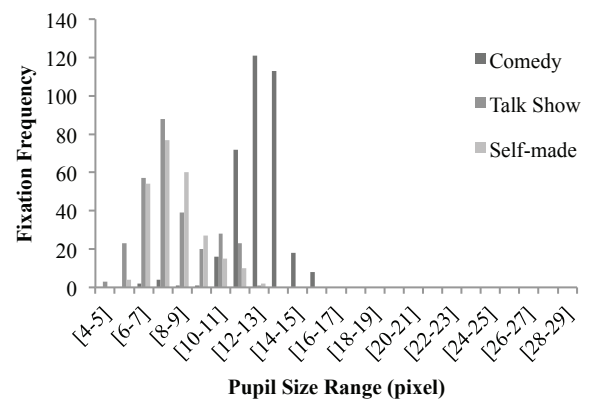


Figure 9 Distributions of Fixation Frequencies per Pupil Size of Subject 2

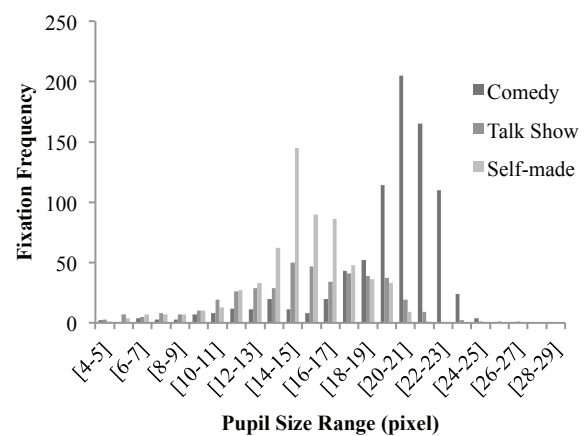


Figure 10 Distributions of Fixation Frequencies per Pupil Size of Subject 3

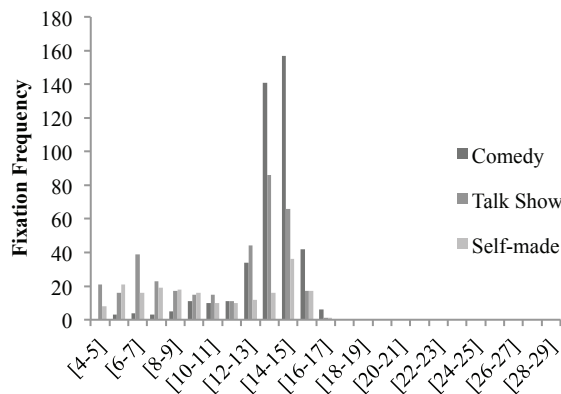


Figure 11 Distributions of Fixation Frequencies per Pupil Size of Subject 4

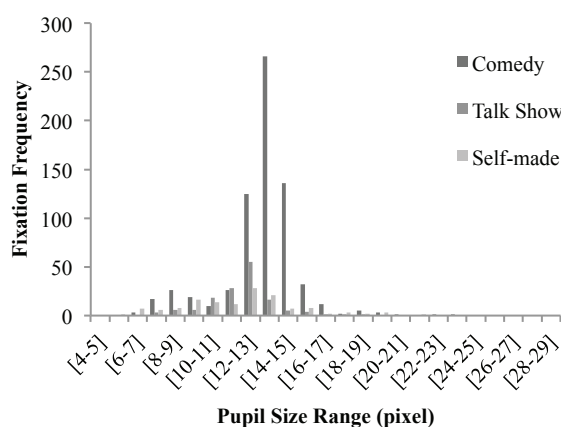


Figure 12 Distributions of Fixation Frequencies per Pupil Size of Subject 5

4. Conclusion

We have presented a study of the relationship among video watching interest, fixation frequency and pupil size. Experimental results show that the high fixation frequency and the large pupil size can be observed in case of high rated video. That shows the possibility of utilizing these two features for analyzing and recognizing video watching interest level objectively. As this study is still in the preliminary stage, further studies including more experiments and more detail analysis are required to be done in the future in order to support our proposed features.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 25330136.

Reference

- 1)Just, M.A. and Carpenter P.A., "Eye Fixations and Cognitive Processes", *Cognitive Psychology*, 8, pp.441-480, October 1976.
- 2)Carniglia, E., Caputi, M., Manfredi, V., Zambarbieri, D. & Pessa E.

"The influence of emotional picture thematic content on exploratory eye movements", *Journal of Eye Movement Research*, 5(4):4, pp.1-9, August 2012.

- 3) Nummenmaa, L., Hyönä, J., & Calvo, M. G., "Eye movement assessment of selective attentional capture by emotional pictures", *Emotio*, Vol.6, No.2, pp.257-268, May 2006.
- 4) Babiker, A.; Faye, I.; Malik, A., "Pupillary behavior in positive and negative emotions", *Signal and Image Processing Applications (ICSIPA)*, 2013 IEEE International Conference on, pp.379 - 383, October 2013.
- 5)R.B. Goldstein, E. Peli, S. Lerner, G. Luo, "Eye Movements While Watching a Video: Comparisons Across Viewer Groups", *Journal of Vision-J VISION* 01/2010, 4(8):643-643, January 2010.
- 6)K. Ong, Y. Ono and W. Kameyama, "Video Summarization Based on Pupil Size, Gazing Point, Heart Rate and Its Evaluation", *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences (Japanese Edition)*, Vol.J93-A, No.11, pp.697-707, November 2010.
- 7)Wei-Ting Peng, Wei-Ta Chu, Chia-Han Chang, Chien-Nan Chou, Wei-Jia Huang, Wen-Yan Chang, and Yi-Ping Hung. "Edit by Viewing: Automatic Home Video Summarization by Viewing Behavior Analysis", *IEEE Transactions on Multimedia*, Vol. 13, No. 3, pp.539-550, June 2011
- 8)Wei-Ting Peng, Wei-Jia Huang, Wei-Ta Chu, Chien-Nan Chou, Wen-Yan Chang, Chia-Han Chang, Yi-Ping Hung. "A User Experience Model for Home Video Summarization", In proceeding of *Advances in Multimedia Modeling*, 15th International Multimedia Modeling Conference, pp.484-495, January 2009.
- 9)Mirametrix S2: <http://www.mirametrix.com/products/eye-tracker/> (Retrieved: 4 August 2014)
- 10) Ogama: <http://www.ogama.net/> (Retrieved: 4 August 2014)