UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Reconstruction of Time-Varying Appeal Inputs that Induce Blink Rate Synchrony

Permalink

https://escholarship.org/uc/item/8j82w0t9

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Author

Nomura, Ryota

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Reconstruction of Time-Varying Appeal Inputs that Induce Blink Rate Synchrony

Ryota Nomura (nomuraryota@waseda.jp) & Go Sato
Faculty of Human Sciences, Waseda University
2-579-15, Mikajima, Tokorozawa City, Saitama, Japan
Go Sato
School of Human Sciences, Waseda University
2-579-15, Mikajima, Tokorozawa City, Saitama, Japan

Abstract

The time-varying appeal of an audiovisual stimulus cannot be directly observed because it is not only determined by the expression itself but also involves the viewer's information processing. In this study, we attempted to reconstruct a time-varying common input that is an appeal to induce blink rate synchrony, i.e., blink rate is suppressed at appealing scenes. In the experiment, 44 (22 male and 22 female) university students watched two videos promoting a local area in Japan while detecting blinks using an eye-tracking device. The results showed that the reconstruction ability was less dependent on the embedding parameters. The results showed that the peak of the reconstructed common inputs did not always correspond to the most impressive scene. In the future, it would be beneficial to apply this method to physiological index data according to the type of attractiveness.

Keywords: blink synchrony; appeal power; nonlinear dynamical systems

Introduction

The purpose of this study is to identify the strength of the appeal of the videos at each point in time that causes the viewer to synchronize blinking based only on the blink rate time series information with considering spontaneous blinking (hereafter simply referred to as *blinking*) as an output of the viewer's cognitive system.

Humans blink approximately 20 times per minute, which is significantly more than the four times per minute required to prevent drying of the ocular surface [1]. Such a high blink rate is thought to be associated with visual attention; Blinking during video viewing is adaptively controlled by the viewer's cognitive processes [2]. Therefore, the flow of cognitive processing is guided by watching the same video and trying to understand the meaning, resulting in similar blink timing across multiple viewers [3].

Visual works such as commercials and videos promoting local areas are created to convey messages to the viewer to buy items or visit the areas. In this context, the term "eye-catching" is a figurative expression that refers to the way an object attracts people's attention and arouses their interest and concern. From the perspectives of cognitive science, however, "eye-catching" may mean, as the term implies, that people look carefully without blinking. However, it is difficult to believe that a few exposures for a short period of time, as in the experiment, would have that much influence. Here, in a more practical sense, we focus on the power of appeal in a limited sense, which refers to the power of a video to induce the blinking behavior of the viewer and to synchronize the increase or decrease of the blink rate, which is the frequency per time.

This attractiveness is not only determined by the expression itself, but also involves the information processing of the viewer, which cannot be directly observed. In recent years, a method has been developed to reconstruct common inputs applied to multiple systems from information about their outputs, which have individual differences. The basic idea of the method, called the Superposed Recurrence Plot (SRP) method [4], is to extract the common dynamics contained in the outputs based on the theory of nonlinear dynamical systems. The output of a single system receiving inputs reflects both the dynamics of the common input and the dynamics of the system. Therefore, the outputs of multiple systems with individual differences commonly contain the dynamics of the inputs in addition to the unique dynamics of each system. Therefore, the state of the common input can be extracted by SRP. By rearranging them on the time axis using the classical multidimensional scaling method, the amplitude of the time series of the inputs at each time point can be obtained [4]. Thus, in this study, we also apply this reconstruction method to the blink rate calculated from the point process data of the blink initiation time of the viewer and attempt to reconstruct the input that increases (decreases) blink rate, which corresponds to the appeal power that cannot be observed directly.

Methods

Participants

The study included 44 participants (22 women) and the mean age was 22.75 years, and the standard deviation was 6.74 years. Recruitment for participation in the experiment was conducted through a portal site that can be viewed by students affiliated with Waseda University. The requirements for participation were those who could stare at a computer screen for about 15 minutes without problems that would interfere with their daily life and those who agreed to have their pictures taken during the experiment. The participants in the experiment were university students and graduate students, including working students.

Stimulation

In this study, we used two videos: the first and second episodes of the "Komoro ga Atsu-i Series," a local PR video for Komoro City, Nagano Prefecture, as visual stimuli. The first installment of the "Komoro ga Atsu-i Series" was

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

released on YouTube in 2016. The local PR video, produced by Komoro City staff with a budget of 9,500 yen, has been picked up by mass media, and as of February 2025, it has been viewed more than 68,000 times. Like the first, the second installment was produced by Komoro City staff and released in June 2017. As of February 2025, it has been viewed over 10,000 times.

Blink measurement

To measure blinks while watching the videos, we video recorded the upper body of the participant, including their face, and visually measured blinks using ELAN6.1. The video of the upper body while watching the video was taken using an iPhone 13 camera placed in front of the monitor using a smartphone stand. To identify the timing at which the local PR video in the video begins playing, we used an application that allows simultaneous shooting (picture-in-picture) with the back and front cameras of the iPhone 13.

The criteria for visually judging blinks were based on research on rapid continuous blinking conducted by Ref. [5] In the Ref. [5], the blinking defined as the pupil being hidden by the eyelids. Therefore, the point at which the eyelid covered the pupil was defined as the starting point of the blink, and the point at which the entire pupil that had been hidden by the eyelid became visible again was defined as the ending point of the blink.

Based on the above criteria, the starting and ending points of the blink were recorded using the annotation insertion function of ELAN6.1. The records were output in a tab-delimited document file format, and the data at the time the blink occurred and the time the blink ended in the video was obtained.

Questionnaire

In the questionnaire survey, new questions were added to the questionnaire used in the research by Ref. [6] on predicting favorability towards local PR videos based on viewing behavior, which measures the level of subjective interest in viewed local PR videos. The questionnaire of Ref. [6] covers the domains of locality, impact, quality, and marketing, based on the video evaluation criteria of the "Gurokaru CM Research Institute," which is operated by Thames Co., Ltd. and researches and introduces local commercials and PR videos originating from local areas. In this study, to identify the scenes that aroused the viewer's interest in the video they watched and the scenes where they lost interest, we added a question to the questionnaire about "scenes that left an impression/distracted the participant" and the reason, a multiple-choice memory question of how much the participant remembered about the information introduced in the video, and a question asking whether the participants knew about Komoro City, Nagano Prefecture before participating in the experiment.

Procedures for video viewing experiment

For each participant, we presented two videos, the first and second installments of the "Komoro ga Atsu-i Series," (https://www.youtube.com/watch?v=99DIBvKWbz0 and https://www.youtube.com/watch?v=U-OmFI-tHVI) which is a local PR video for Komoro City, Nagano Prefecture, using the following procedure (**Fig. 1**.).

Prior to the video viewing experiment, participants were asked to consent to the content of the experiment, the acquisition of audio and images, and the handling of personal information. After consent was obtained, the participants were instructed to watch the video in the viewing position. Thereafter, when measuring the gaze data using the measurement device, we performed calibration to confirm that the line of sight was being measured correctly. After completing the calibration, we placed the iPhone 13 on a smartphone stand placed in front of the video playback display and started recording video. Five seconds after the commencement of video recording, the first video was presented. The presentation order of the two videos was changed for each participant to eliminate order effects. Additionally, in Ref. [7], Takano stated that there were multiple cases where participants got excited after seeing the title or thumbnail of a video and play it, but then "stopped watching halfway because the content is not interesting." Therefore, to approximate the actual viewing situation on YouTube, we presented a thumbnail image of the video for 15 seconds before presenting the video. After watching the video, we conducted a questionnaire survey. The participants watched the second video and measured and collected each data in the same way. After viewing the two videos and completing the questionnaire survey, we conducted a verbal survey of the participants' eye preference to use it as a reference for analyzing the gaze data, and the experiment

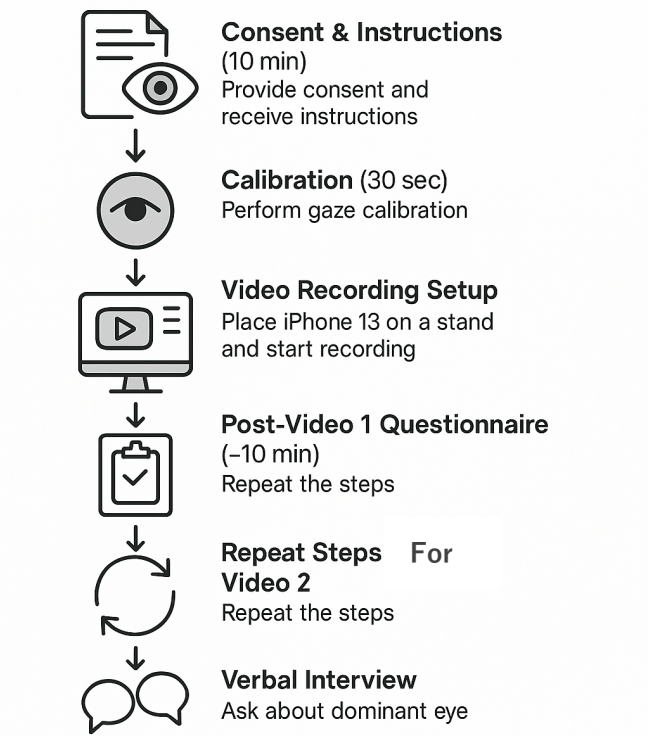


Figure 1 Experimental procedure for video viewing study.

ended. We conducted an experiment using this procedure on 44 people.

It took about 10 minutes to explain the experiment and confirm consent, about 30 seconds to calibrate, about 4 minutes to watch the video, and about 10 minutes to answer the questionnaire. The total experiment time per person was approximately 30 minutes.

All procedures of this study have been approved by the research ethics committee of Waseda University (2021-311).

Reconstruction methods

We calculated the blink rate using the point process time series of the blink onset by shifting a 5-second time window allowing 1 second overlap. We then applied the SRP method to the blink rate time series [4]. First, we transformed the time series data of blink rate into a state space using a time-delay coordinate system, and calculated recurrence plots (hereafter, RPs) for each participant. Next, we identified the common components of the RPs as follows: We defined the (i,j)-th element $\boldsymbol{SR}_{\boldsymbol{ij}}^{\boldsymbol{q}}$ of the superposed recurrence matrix as

$$\boldsymbol{SR}_{\boldsymbol{ij}}^{\boldsymbol{q}} = U\left\{\left(\sum_{k=1}^{N} \boldsymbol{R}_{ij}^{k}(\varepsilon_k)\right) - q\right\}, i,j = 1,2,\dots,N, \quad (1)$$

where N is the total number of the blink rate time series data and k is for each participant in the experiment. In Eq (1), $\boldsymbol{R}_{ij}^{k}(\varepsilon_k)$ is the (i,j)-th element of the recurrence matrix at threshold ε_k. The term $U(\cdot)$ is a unit step function that returns 1 if the argument is greater than or equal to 0; and returns 0 if the argument is less than 0. In other words, Eq. (1) determines if a greater number of $\boldsymbol{R}_{ij}^{k}(\varepsilon_k)$ are equal to 1 compared to the number of superposition q.This means that the (i,j)-th element on $\boldsymbol{SR}_{\boldsymbol{ij}}^{\boldsymbol{q}} = 1$ if $\boldsymbol{R}_{ij}^{k}(\varepsilon_k) = 1$ in at least q systems, otherwise $\boldsymbol{SR}_{\boldsymbol{ij}}^{\boldsymbol{q}} = 0$. For example, for $q = 3$, the (i,j)-th element of $\boldsymbol{SR}_{\boldsymbol{ij}}^{\boldsymbol{q}}$ will not be 1 unless $\boldsymbol{R}_{ij}^{k}(\varepsilon_k) = 1$ for at least 3 systems. The states of that (i,j)-th elements for which $\boldsymbol{R}_{ij}^{k}(\varepsilon_k) = 1$ are neighbors in the dynamics of the common input and are also neighbors in the dynamics of the individual systems. Therefore, if the state at a given time is a neighbor in at least q systems, it is likely that the state at that time is also a neighbor in the dynamics of the common input. Note that equation (1) satisfies the formulation of the SRP method [4] in the previous study when $q = 1$. By rearranging them on the time axis using the classical multidimensional scaling method, the amplitude of the time series of the inputs at each time point can be obtained [4]. In this method, the distance between the states at each time point is strictly defined in the target state space, satisfying the assumptions of the multidimensional scaling construction method.

The dimension of the blink rate is not known, but the complexity of the biological signal suggests that it is not of the high dimension assumed. Therefore, the embedding parameters for the time-delay coordinate system were determined based on a previous study [5]. Namely, the embedding dimensions were set to $m = 3, 4, 5$ and the time delay $\tau = 1$ (1 second). The number of superpositions $q = 1, 2, 3$.

The units of the reconstructed common input are arbitrary due to the loss of scale information. The direction of the positive and negative values is also arbitrary, but in this study the direction in which the number of blinks increases was defined as positive for ease of interpretation.

Other analyses

The analysis of the number of blinks used data from 41 participants in the first video and 43 participants in the second video. Other data were omitted from the analysis because the video could not be fully recorded during the video viewing experiment due to equipment failure.

To analyze the relationship between the number of blinks and the interest-stimulating effect of local PR videos, we first created a time series of the times at which blinks occurred. We then examined whether there was a difference in the occurrence of blinks between scenes that were considered memorable and other scenes.

Results

Parameter dependence of reconstructed common input

The common input could be reconstructed for the parameters considered in this study. First, for the first video, when the number of superposition $q = 1$, the amplitude of the reconstructed common input was small; compared to when $q > 1$, it was flat and did not adequately capture the variability. The trend for $q = 1$ was similar for the second video.

In contrast, for $q > 1$, the peaks of the reconstructed common input did not depend much on the parameters (**Fig. 2**(a), (c)). When the number of superpositions was fixed and the embedding dimension was considered (**Fig. 2**(c)), there was no significant difference in the amplitude of the reconstructed common input for $m > 3$ ($r \approx .90$). This trend was also common in the second video. According to the Ref. [4], the peaks can be adequately detected even when the method is applied to the output of systems that exhibit chaotic responses. Therefore, the peaks found here are considered to reflect the power of the video.

Relationship between the reconstructed common input and the number of blinks

In the reconstructed common inputs, three peaks were found in the first video. Two peaks were found in the second video. In contrast to the number of blinks (average value per 60 seconds), which fluctuated over a short period of time in the order of 10^0 seconds, the reconstructed common input showed slow fluctuations in the order of 10^1 seconds.

For the first video, the first peak of the reconstructed common input increased slightly before the increase in the number of blinks. This corresponds to just after one of the memorable scenes (9[s]–17[s]). The two peaks after the reconstructed common input also showed an increase in the

number of blinks, but their values were similar to those seen in other scenes.

For the second video, there was no clear correspondence between the reconstructed peak of the common input and the increase in blink rate. The reconstructed common input also did not correspond to the scenes that were memorized.

Discussion

In this study, we attempted to reconstruct the time-varying appeal, which cannot be directly observed, as a common input based solely on information from the viewer's blink rates. As the result, we found that the common input varied more slowly than the variation of the blink rate. In particular, when $m > 3$, the general shape of the reconstructed common input did not change. Nonlinear dynamics theory suggests that embedding is possible when sufficient dimensionality is ensured, and in this case $m = 4$ would be the sufficient dimensionality.

When the state recurred at the same time for multiple participants, corresponding to $q > 1$, there were scenes in which the blink rate increased after the common input for the direction of increasing blinks. In contrast, for the direction of decreasing blink rate, there was no clear correspondence between the common input and the scene in which the blink rate decreased. Consequently, the reconstructed common input peaks showed little correspondence with the scenes reported to be memorable. However, the lack of correspondence was also true for the number of blinks. Therefore, it is possible that the intensity of postmortem memory for memorable scenes does not necessarily correspond to the blinking behavior associated with sequential information processing.

One reason for these results may lie in the method used to measure impressionability. In the experimental procedure of this study, the participants were asked to respond to the most memorable scenes by writing freely, so the strength of impressions of scenes other than those mentioned was not clear. Therefore, in future research, it is possible to clarify the statistical relationship between the reconstructed common input in the order of seconds by incorporating a method in which the participants are asked to continuously evaluate subjectively by using a device such as an emotion evaluation dial [8].

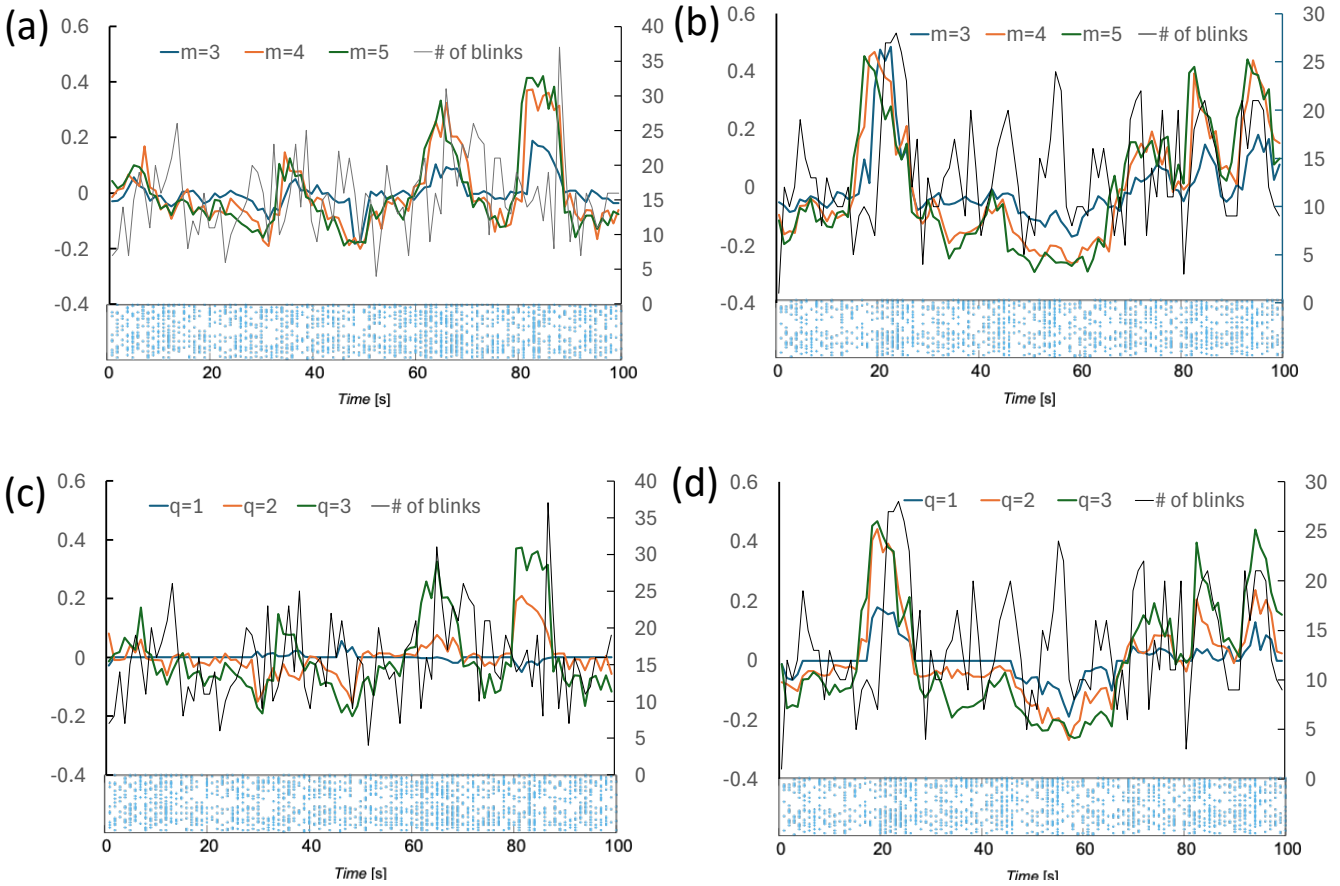


Figure 2 Reconstructed common inputs and raster plots of blink for the first (left) and second (right) videos. (a), (c) for $m = 4$, $q = 1, 2, 3$ and (b), (d) for $m = 3, 4, 5$, $q = 3$. The light blue dots indicate the blink onset point process (bottom of each subfigure). The number of blinks (right vertical axis of each subfigure) is that the value per 60 seconds averaged over the participants. The length of the entire video is plotted as 100 so that the results of the two videos can be compared.

The correlation between the reconstructed common input and the blink count time series was not large. If this result means that the common input was adequately reconstructed while the output blink rates were diverse due to individual differences in the viewer's cognitive system, it would be useful to investigate whether this result can be replicated in future research by simultaneously feeding the reconstructed common input to a blink model [2] that incorporates individual differences.

The SRP method in this study can be widely applied to non-stationary time series data because it does not make any statistical assumptions such as a specific normal distribution in the background of the analysis [4]. Therefore, in addition to blink, it may be possible to directly approach the appeal by applying the SRP method to other physiological index data such as EEG, heart rate, body surface temperature, skin potential, and respiration. Since the common input is reconstructed as a time series, the SRP method is expected to have a wide range of applications, including not only the study of synchrony [8] caused by the common input as an appeal power, but also the prediction of influence and causal estimation [9].

References

[1] Nakano, T., Yamamoto, Y., Kitajo, K., Takahashi, T., & Kitazawa, S., (2009). "Synchronization of spontaneous eyeblinks while viewing video stories", *Proceedings of the Royal Society B: Biological Sciences*, 276(1673), 3635–3644.

[2] Hoppe, D., Helfmann, S., & Rothkopf, C. A., (2018). "Humans quickly learn to blink strategically in response to environmental task demands", *Proceedings of the National Academy of Sciences*, 115(9), 2246–2251.

[3] Nomura, R. and Okada, T. (2014). "Spontaneous Synchronization of Eyeblinks during Story-Telling Performance", *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, 21(2), 226–244.

[4] Nomura, R., Fujiwara, K., & Ikeguchi, T., (2022). "Superposed recurrence plots for reconstructing a common input applied to neurons", *Physical Review E*, 106(3), 034205.

[5] Nomura, R., & Maruno, S. (2019). Rapid serial blinks: An index of temporally increased cognitive load. *PLOS ONE*, 14(12), e0225897.

[6] Murata, H. (2022). Predicting favorability for local PR videos based on viewing behavior. Graduate Thesis submitted to School of Human Sciences, Waseda University (unpbulished).

[7] Takano, Y. (2018). *Local CMs that Sell Things Local PR Videos that Gain Things Local PR Videos that Win the Hearts and Minds of All of Japan, Marketing Strategy*, Gentosha.

[8] Ruef, A. M., & Levenson, R. W., (2007). "Continuous measurement of emotion", In *Handbook of Emotion Elicitation and Assessment*, 286–297.

[9] Nomura, R. (2024). "Reliability for music-induced heart rate synchronization", *Scientific Reports*, 14(1), 12200.

Acknowledgement

We were able to conduct this research with the cooperation of Komoro City, who kindly agreed to allow us to use their local PR video. The viewing experiment using the unique and heartwarming works of Komoro City was an invaluable experience for us to rediscover the charms of Komoro City and the passion of the local community. We would like to express our gratitude to everyone involved for their generous cooperation.

Note

The data of this study was also used in the second author's graduation thesis (unpublished) submitted to the Faculty of Human Sciences, Waseda University, in the school year 2023. However, the analysis was originally performed, and the method and other texts were added and modified to the thesis. Additionally, this study was originally scheduled to be presented at the 41st Annual Meeting of the Japanese Cognitive Science Society but was withdrawn due to scheduling conflicts. Hence, this is the original presentation. Author Contribution: The first author performed the analyses and wrote an original manuscript; The second author designed and performed the experiments.