Relating aesthetic-value judgment to perception: An eye-tracking and computational study of Japanese art *Ukiyo-e*

Yuka Nojo (nojoyuka@g.ecc.u-tokyo.ac.jp)[1]
Graduate School of Arts and Sciences, The University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo, 153-8902, Japan

Tomoyuki Maekawa (maekawa@ailab.ics.keio.ac.jp)[2]
Graduate School of Science and Technology, Keio University
3-14-1 Hiyoshi, Kohoku-Ku, Yokohama-shi, Kanagawa, 223-8522, Japan

Yuri Sato (sato.yuri@ocha.ac.jp)[3]
Faculty of Core Research, Ochanomizu University
2-1-1 Ohtsuka, Bunkyo-ku, Tokyo, 112-8610, Japan

Kazuhiro Ueda (ueda@g.ecc.u-tokyo.ac.jp)[1]
Graduate School of Arts and Sciences, The University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo, 153-8902, Japan

Abstract

Aesthetic value, beauty, is a complex concept in that it has both subjective and objective aspects. However, in previous eye-tracking studies of artworks, the target associated with the gaze has often been only the latter image feature (e.g., symmetry). By contrast, recent developments in computational aesthetics (especially aesthetic classifiers) offers a path that treats these two aspects comprehensively. Along this line, we further develop eye-tracking research. We conducted computational model-based data analyses of eye-tracking behaviour, with the aim of providing more fine-grained insights into the comprehensive concept of beauty. In contrast to previous studies that focused on Western fine-artworks, our study is distinctive in that it focuses on visual art in Eastern culture, namely Japanese *ukiyo-e* (the well-known works of Katsushika Hokusai and Utagawa Hiroshige). Our empirical results showed that the gaze trajectory and movement area differed significantly between highly-rated and non-highly-rated paintings. This provides positive evidence for the relationship between aesthetic value and perceptual behaviour.

Keywords: experimental aesthetics; computational aesthetics; eye-tracking; gaze pattern; aesthetic value; ukiyo-e

Introduction

This study aims to address the question of whether or not human judgments related to value, especially those related to aesthetic value, can be objective; in other words, can such judgments be substantiated? Unlike judgments involving logic and mathematics, these questions are not self-evident. Although this type of discussion about aesthetics has its origins in the discourses of Plato and Aristotle, Sibley (1965) is the one who formulated the question from a modern perspective. Based on his linguistic analysis (analytic philosophical) of beauty, Sibley argued that the concept of beauty is divided into the aspects related to the subjective evaluation of the viewer (expressed in words such as good, bad, excellent, and mediocre) and those related to the descriptive attributes of the target image (expressed in words such as graceful, garish, and balanced). See also Sibley (2001) for a comprehensive collection of his papers. Birkhoff (1933) and empirical (behavioural) research have focused mainly on the latter aspect of beauty. Typical examples are studies measuring the viewers' gaze such as those by Khalighy, Green, Scheepers & Whittet (2015), Lacoste Badiea, Gagnan & Droulers (2020), Nojo (2015), Lüttge & Souza (2019), Lakhal, Darmon, Bouchaud & Benzaquen (2020), and Christensen, Ball & Reberc (2020), where symmetry, golden ratio, and/or complexity have been addressed. In contrast, this study examines how the complex concept of aesthetic value or beauty that mixes subjectivity and objectivity can be empirically addressed.

One of the promising approaches to this problem is computational aesthetics that builds aesthetic classifiers based on the framework of statistical machine learning (cf., Galanter, 2012; Spratt & Elgammal, 2015). This approach can effectively link human evaluations of aesthetic value, whether subjective or objective, to the features on the image being evaluated. Specific examples include the aesthetic evaluation judgments of photographs using decision trees (Datta, Joshi, Li & Wang, 2006), cluster classification (Hayn-Leichsenring, Lehmann & Redies, 2017), and naive Bayesian classifiers (Li & Chen, 2009). We aim to further develop research in this direction by adding comprehensive aesthetic-value ratings, image-level data, and eye-tracking data.

Behavioural scientific studies on eye tracking and gaze also have the advantage of incorporating computational models into gaze data analysis, allowing for a more detailed analysis. While previous studies mainly focused on analysing where fixation points are located (Ganczarek, Pietras, Stolinska & Szubielska, 2022; Locher, Krupinski, Mello-Thoms & Nodine, 2007), our study uses hidden Markov model (HMM) (cf., Baum & Petrie,1996; Chuk, Crookes, Hayward, Chan & Hsiao, 2017) to analyse gaze data, which enables a more fine-grained gaze pattern analysis (movement area, movement trajectory, etc.).

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

Task analysis

This study focuses on *ukiyo-e* as a pictorial art form. It is a painting genre that developed mainly in Japan from the 17th to 19th centuries and is known to have influenced Europe (especially French Impressionism) in the 19th century (cf., Bell, 2004; Harris, 2012). As experimental stimuli, we used Katsushika Hokusai's 'Thirty-six Views of Mount Fuji' (1831-34) and Utagawa Hiroshige's 'Fifty-three Stations of the Tokaido Road' (circa 1834), known in art research as a group of landscape paintings representing *ukiyo-e* (woodblock prints) (cf., Forrer, 2017; Nagata, 1995; Uhlenbeck & Jansen, 2008). Most studies on aesthetic classifiers have focused on perspective-based images (using traditional techniques of Western painting). However, *ukiyo-e* differs from traditional Western techniques in several ways; it does not fully employ perspective (Kadar & Effken, 2008), and uses a limited colours palette, among other differences. Therefore, incorporating the beauty of *ukiyo-e* into the image characteristics obtained in previous studies may be difficult. These aspects of *ukiyo-e* are distinctive in art paintings and are targets for empirical studies.

The relationship between emotion and gaze in viewing images and visual artwork has been studied to some extent. For instance, eye movements have been reported to change depending on whether images elicit positive (e.g., happy) or negative (e.g., sad) emotions. Yang, Sunaga, Toh and Ihara (2018) and Tarnowski (2020) reported on the duration time of gaze; Locher et al. (2007) and Yanulevskaya, Uijlings, Bruni, Sartori, Zamboni, Bacci, Melcher & Sebe (2012) analysed the fixated region of gaze. Locher et al. (2007), which used landscapes, reported no significant correlation, whereas Yanulevskaya et al. (2012), which used abstract paintings, reported more fixated regions when positive emotion was evoked.

Interestingly, some artists, reflecting on their own activities, claim the following (Hamm, 1982). A masterpiece of landscape painting has the characteristic that has an introductory point to guide the viewer's gaze: while following motifs such as trees and rocks, the viewer gradually moves the gaze from the introductory point to the back of the painting; the viewer then sees the whole painting without focusing on any particular point; afterwards the viewer returns the gaze to the introductory point, from the back towards the front; the viewer again moves the gaze to follow the motif and so on. This game-like mechanism guides the viewers' gaze and allows them to understand the context of the painting while moving smoothly through a picture.

The characteristic which Hamm suggests, i.e., that the gaze moves throughout the painting in a good landscape painting, can be interpreted as meaning that the larger the area over which the gaze moves, the greater the landscape painting is: in this study, we will test whether this is true. We expect to find that the area of the viewer's eye movement in *ukiyo-e* landscape paintings, is larger for those that are rated high in terms of "beauty" (a composite concept that is both subjective and objective) than for those that are rated low, and that certain characteristics can be extracted from the trajectory of the viewer's eye movement. To examine this, we conducted an experiment using eye-tracking.

The structure of the paper is as follows. In the next section, a pilot experiment was conducted to identify the necessary experimental setup. The method of the main experiment including the data analysis method was then described. After that, the results of data-analyses were explained. Finally, general discussions based on these empirical findings were provided.

Pilot experiment

Participants Twenty-seven university students at the University of Tokyo (15 males, 12 females, *mean age*=20.2 years, *SD*=2.31) participated in the experiment. None of the participants were art specialists. The participants were paid a fixed amount as compensation after the experiment.

Stimuli To evaluate viewer impressions of landscape paintings and obtain gaze data, we selected 46 paintings of 'Thirty-six Views of Mount Fuji' by Katsushika Hokusai (1760-1849), who is known for his "cartoonish and illustrative style" (cf., Ando, 1859; Tsuji, 1972); and 55 works from 'Fifty-three Stations of the Tokaido Road' by Utagawa Hiroshige (1797-1858), known for his "realistic style" (cf., Ando, 1859; Tsuji, 1972). All the images were obtained from Wikipedia[1]. Although the four corners of 'Fifty-three Stations of the Tokaido Road' have decorative window-like frames, all frames were erased to ensure that the experimental stimuli for the viewers remained constant, in terms of shape, between the two painters. Accordingly, the viewers were presented with paintings that had nothing drawn in the four corners.

Procedure They viewed paintings randomly presented on the display individually for 30 seconds, and then answered a question asking them to rate their impressions of the paintings. After viewing each painting, participants were asked to rate on a 10-point scale (1 = most negative, 10 = most positive) the "quality of composition", "integrity of colour harmony", "static or dynamic", "favourability", and "beauty" (see Carbon, 2017).They then answered in a Yes/No format whether they had seen the painting before (see also Smith, Smith & Jeffery, 2017). Viewing time was measured in the same way as in previous studies (cf., Carbon, 2017; Smith et al., 2017; Smith & Smith, 2001). The gazes of the participants while viewing the paintings were measured using a Tobii TX300 Pro eye-mark recorder. Viewing the paintings and responses to the questions were considered as one trial, and data from 29 trials per participant were collected. Based on these findings, we designed the following setup for the main experiment.

[1] Wikipedia; Thirty-six Views of Mount Fuji (version:15 January 2019) https://en.wikipedia.org/wiki/Thirtysix_Views_of_Mount_Fuji, Wikipedia; The Fifty-three Stations of the Tōkaidō (version: 15 January2019) https://en.wikipedia.org/wiki/The_Fifty-three_Stations_of_the_T%C5%8Dkaid%C5%8D

Table 1: Impression rate of the pilot experiment

		Pilot Experiment	
		Beauty + Favorability	
		Color	
		Average	SD
High	Shinagawa: Sunrise	9.20	1.07
	Under the Wave off Kanagawa	8.00	0.90
	Ejiri in Suruga Province	7.65	2.06
	Under the Mannenbashi Bridge at Fukagawa	7.50	1.30
	Viewing Sunset over the Ryogokubashi Bridge from the Ommayagashi River Bank	7.50	1.46
	Okazaki:Yahagi Bridge	7.30	1.99
Low	Futagawa: Sarugababa	3.35	1.18
	Fujieda: Changing Porters and Horses	3.80	2.26
	Okitsu: The Okitsu River	4.65	2.86
	Akasaka: Inn with Serving Maids	4.70	0.75
	Sea Lane off Kazusa Province	4.75	2.79
	Fujikawa: Scene at the Boundary Marker	4.85	2.43

Result Table1 shows the names of the twelve pictures including six paintings with particularly high impression ratings of "beauty" and "favourability" (*mean*=7.85, *SD*=0.70) and six paintings with low impression ratings (*mean*= 4.34, *SD*=0.62), resulting from the pilot experiment (*mean*= 6.17, *SD*=2.17). The highest-rated pictures were 'Under the Wave off Kanagawa' (commonly known as 'The Great Wave off Kanagawa'), 'Shinagawa: Sunrise', 'Ejiri in Suruga Province', 'Under the Mannenbashi Bridge at Fukagawa', 'Viewing Sunset over the Ryogokubashi Bridge from the Ommayagashi River Bank' and 'Okazaki:Yahagi Bridge'; the lowest-rated pictures were 'Futagawa: Sarugababa', 'Akasaka: Inn with Serving Maids', 'Sea Lane off Kazusa Province', 'Fujikawa: Scene at the Boundary Marker', 'Okitsu: The Okitsu River' and 'Fujieda: Changing Porters and Horses'

Main Experiment

Participants Forty-eight university students (30 males and 18 females, *mean age*=22.2 years, *SD*=3.37) recruited from within and outside the University of Tokyo participated in the experiment. None of the participants were art specialists. The participants were paid a fixed amount as compensation after the experiment.

Stimuli We used image data from museum collections (Source of image data: The Sumida Hokusai Museum, Ota Memorial Museum of Art) and excluded paintings whose colours were significantly different from those in the pilot experiment. A total of 24 images were used in the main experiment: 12 original images and 12 black-and-white versions to examine colour effects: Six of the original images were rated particularly high and the remaining six were rated particularly low in the pilot experiment.

Procedure The same procedure as in the pilot experiment was used to obtain data for 24 paintings per participant in the experiment; each run of viewing the paintings and answering the questions were considered as one trial.

Data-Analysis Method

Total area of eye movement in 30 seconds Data from 46 individuals were used in the analysis, excluding two individuals whose fixations were less than 34% of the mean number of fixations in the 48 individuals. To focus on the relationship between impression ratings and the area of eye movement between different pictures, we calculated the impression ratings and the total area of eye movement for each picture in the following manner, regardless of the participant. The impression rating was the sum of the "favourability" and "beauty" ratings, and the average of these sums was calculated for each image. We checked whether the impression ratings of "favourability" and "beauty" significantly differed between the two groups of 12 original pictures, of which six were rated highly in the pilot experiment and six were rated low. For the total area of eye movement, we calculated the index of eye movement over a wide area of the work during viewing, referring to the method of Tsujimoto & Murakami (2010). Fixation was determined by an eye-mark recorder. The ratio of the area of the polygonal region to that of the painting was averaged for each painting to obtain the total eye-movement area.

Gaze pattern analysis In order to see whether gaze patterns differ significantly between highly and low rated paintings, we extracted gaze patterns using HMMs (Chuk et al., 2017). HMMs are statistical models, and their results can be observed in a set of stochastic processes. Patterns can be extracted from time-series data such as gaze data, and gaze behaviours can be visualized by learning from eye fixation data. We compared the gaze behaviours between highly and low rated paintings. Similar gaze patterns are aggregated into a joint HMM by a variational hierarchical expectation maximization (VHEM) algorithm (see Ellahi, Vigier & Le Callet, 2020). In the model proposed by Chuk et al. (2017), regions of interest (ROIs) are treated as hidden layers, the transition matrix is the observer's saccade behaviour, and the number of hidden layers (ROIs) is automatically obtained within a maximum number using the variational approach of Bayesian inference.

In this study, we used HMM (Chuk et al., 2017) to extract gaze patterns using 30 seconds of gaze points per painting for each of the 46 participants, and grouped them into one representative HMM with the VHEM algorithm. The maximum number of ROIs was exploratively set to 18, and all images were processed with the same parameters. In this analysis, in order to have the consistency among the images, the number of ROIs is fixed and the initial values are the same for all images. The ROIs are learned from the gaze data. The most likely order of ROIs is calculated from the initial probability and transition matrix. The ROI with the highest prior probability becomes the most likely first fixation. The most likely second fixation among the unvisited areas becomes the next ROI with the highest probability, under the condition of the first fixation. As a result, the code permutes the states so that ROI1 is the most likely first fixation and ROI2 is the most likely second fixation resulting from ROI1. In this way, the chronological order can be roughly visualized

Table 2: Impression rate and total eye movement

		Main Experiment							
		Beauty + Favorability				Total eye movement			
		Color		Grayscare		Color		Grayscare	
		Average	SD	Average	SD	Average	SD	Average	SD
High	Shinagawa: Sunrise	7.53	1.37	5.78	1.66	41.1%	7.6%	45.8%	9.6%
	Under the Wave off Kanagawa	8.26	1.31	7.26	1.64	49.4%	10.5%	49.1%	13.7%
	Ejiri in Suruga Province	7.51	1.61	6.90	1.75	54.4%	8.3%	53.6%	10.0%
	Under the Mannenbashi Bridge at Fukagawa	7.48	1.76	6.20	2.02	51.0%	7.7%	50.1%	8.5%
	Viewing Sunset over the Ryogokubashi Bridge from the Ommayagashi River Bank	6.89	1.84	6.30	1.77	45.0%	12.2%	45.0%	7.1%
	Okazaki:Yahagi Bridge	7.36	1.73	6.39	1.83	40.2%	8.0%	43.7%	6.6%
Low	Futagawa: Sarugababa	6.72	1.76	5.70	1.76	39.2%	8.8%	40.7%	8.2%
	Fujieda: Changing Porters and Horses	5.83	2.21	4.78	1.75	41.5%	6.2%	42.0%	9.9%
	Okitsu: The Okitsu River	6.08	1.73	5.33	1.72	46.4%	9.3%	48.4%	11.2%
	Akasaka: Inn with Serving Maids	6.21	1.76	5.63	1.87	38.8%	7.1%	41.7%	10.4%
	Sea Lane off Kazusa Province	6.45	1.83	5.67	1.76	43.5%	9.0%	47.6%	9.4%
	Fujikawa: Scene at the Boundary Marker	6.13	1.85	5.02	1.77	42.9%	6.1%	44.9%	8.3%

from the ROIs.

Results

Total area of eye movement in 30 seconds

Table 2 shows the names of the six pictures classified in the high rating group and the six pictures classified in the low rating group along with their total favourability and beauty, and total area of eye movement.

To test the difference in the mean total area of eye movement between the high and low rating groups during the 30s for each of the 46 participants, a logit transformation was performed. The results showed a significant difference in the mean total area of eye movement between the high-rating and the low-rating groups ($t(10)=1.849$, $p= 0.047$), supporting our hypothesis. For the grayscale images, a one-sided t-test was conducted using the same procedure as that for the original images, since no significant difference was found in the variance after logit transformation. The result showed a significant difference ($t(10)=1.812$, $p=0.049$).

Figure 1 shows the results of plotting the trajectories of fixations for the original images classified into the high and low rating groups and the gaze data for each image selected one trial at a time. Following the same procedure as in the previous section, no significant difference was found in the variance after logit transformation, a one-sided t test resulted that the total area of gaze movement was not significantly different between the colour and grayscale images ($t(10)=1.71$, $p=0.19$). However, in 10 of the 12 images, the grayscale image had an average +1.8% larger total area of line-of-sight movement than the other images.

The average of the total area of eye movement during 30 seconds for the 12 paintings was 44.5%. This supports the findings of Locher et al. (2007), who revealed that viewers did not see 54% of the total area of the painting even if they viewed the paintings without a time limit.

Gaze pattern

In most of the pictures, fixations tended to be concentrated at the centre of the picture. This tendency supports previous studies (cf., Locher et al., 2007). Exceptions of this pattern are '*Ejiri in Suruga Province'*, '*Under the Mannenbashi Bridge at Fukagawa'* and '*Viewing Sunset over the Ryogokubashi Bridge from the Ommayagashi River Bank'* (all in '*Thirty-six Views of Mount Fuji*'). The typical gaze pattern while viewing these paintings was a repeated back-and-forth movement between the upper left (signature) and the right edge (the person sitting on the bow, the bridge embankment, and the island on the shore) across the small Mt. In these paintings, the gaze moved back and forth between motifs placed far apart, which is believed to have increased the total area of movement of the gaze. The pattern in which the gaze circumnavigated the entire picture plane was observed in the majority of the trials for 'Thirty-six Views of Mount Fuji; Under the Wave off Kanagawa', which received the highest impression rating, and 'Fifty-three Stations of the Tokaido Road; Shinagawa: Sunrise', which received the second highest impression rating. Even though the upper right area of '*Under the Wave off Kanagawa'* depicts only the sky, eye fixations also occurred in that area. This type of eye movement to the area where no motif, such as a signature, was drawn, was observed in the six pictures with high ratings, suggesting the existence of a unique mechanism that induces eye movement.

In contrast, the typical gaze pattern while viewing paintings with low impression ratings was that the participants continued to gaze at the boat and people carrying luggage depicted in the centre of the painting, and their gaze moved less frequently away from the centre of the painting to the spaces above, below, right, and left. This pattern of gazing

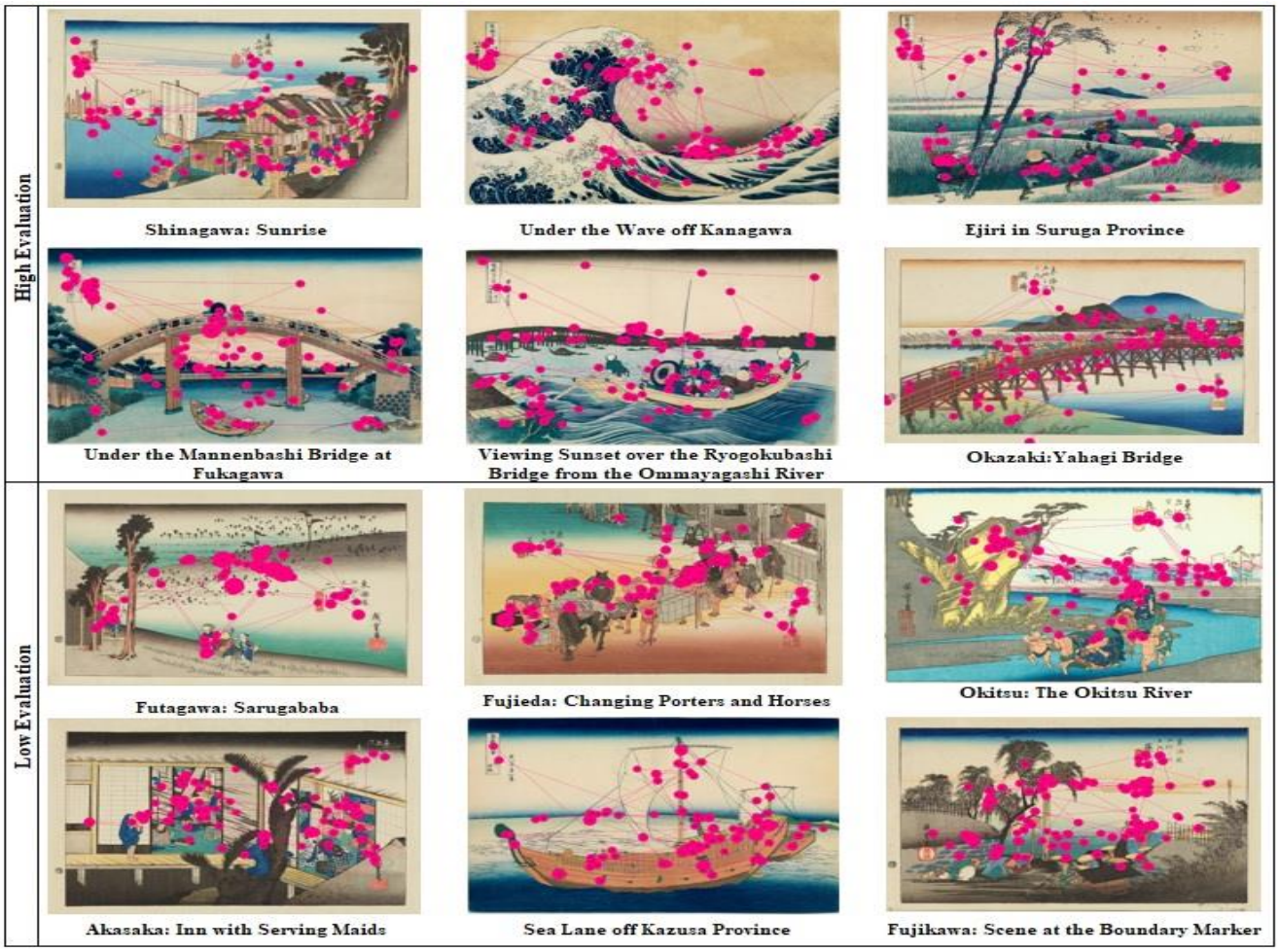
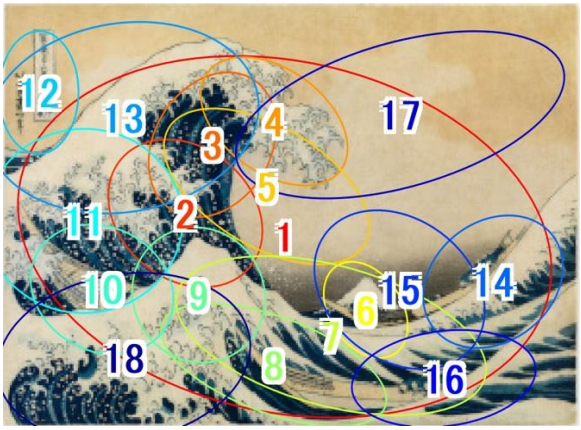


Figure 1: Pictures classified into the high and low rating groups in the pilot experiment and the fixation trajectories in the gaze data for one trial for each picture

was particularly noticeable in '*Fujieda: Changing Porters and Horses*', '*Akasaka: Inn with Serving Maids*', and '*Okitsu: The Okitsu River'* (all '*Fifty-three Stations of the Tokaido Road*'). In addition to the above pattern, '*Sea Lane off Kazusa Province*' which received the lowest impression rating among '*Thirty-six views of Mount Fuji*', the small Mount Fuji in the distance was rarely gazed at. We analysed the raw gaze of 46 people for 12 pictures, each for 30 seconds, using the HMM tool based on a previous study (Chuk et al., 2017). The results show a spiral pattern of gaze turning from the centre to the outside in '*Under the Wave off Kanagawa*', [Fig.2] '*Under the Mannenbashi Bridge at Fukagawa*', '*Okitsu: The Okitsu River*' and '*Fujikawa: Scene at the Boundary Marker*'. The transition information between ROIs was described by the transition matrix [Fig2]. The transition probability between the 18 ROIs were exceeding 97%, which indicates that the viewer's gaze stayed around each ROI for a certain period. In this analysis, the order of the ROIs is time-series, so the transition matrix showed that eyes move from ROI1 to ROI18. In some pictures, the size of first or second ROI, early in the viewing process, are quite large [Fig2]. It means there was no consistent strategy for those fixation behaviours. However, we can still see a trend in terms of how spread out the ROI is, whether it is explorative or focused.

Although the ideal eye movement mentioned in the Task Analysis, "entering from the introductory point, moving toward the back and finally returning to the original position (introductory point)", was not observed; instead "gradually moved from the introduction point to the back of the painting while following motifs" and a spiral-like "gaze moves more widely" behaviour were found. In case of 'Under the Wave off Kanagawa', the viewer's eye follows the movement of the large wave from bottom to top, then follows the motifs in the following order: top of the splash of great waves, Mt.Fuji (which should have the most depth), a small boat floating between the waves in the foreground of Mt.Fuji, a medium wave slightly to the left in the foreground, and a second small boat floating to the left foreground. This movement was from back to front. In all of the highly rated pictures include motifs that evoke a sense of depth (Mt.Fuji, a boat floating offshore, etc.), which overlap with the centre point of the ROI (the point where the distribution of gaze is most concentrated),



ROI from 1 to 13 draw a spiral

	ROI1	ROI2	ROI3	ROI4	ROI5	ROI6	ROI7	ROI8	ROI9	ROI10	ROI11	ROI12	ROI13	ROI14	ROI15	ROI16	ROI17	ROI18
ROI1		11.8%	11.1%	5.7%	9.7%	9.4%	7.2%	5.6%	3.1%	4.0%	1.7%	5.9%	3.2%	6.6%	8.6%	2.1%	3.4%	0.8%
ROI2	30.3%		10.8%	2.9%	9.8%	1.0%	5.6%	2.0%	7.4%	8.8%	9.5%	3.1%	3.8%	0.6%	1.0%	0.4%	1.0%	2.0%
ROI3	36.0%	16.3%		15.9%	10.7%	0.4%	0.8%	0.4%	0.4%	0.7%	2.6%	1.1%	11.5%	0.5%	0.6%	0.4%	1.4%	0.4%
ROI4	37.9%	3.5%	19.3%		20.5%	0.5%	0.5%	0.5%	0.9%	0.6%	0.4%	1.0%	4.6%	0.8%	2.0%	0.5%	6.5%	0.2%
ROI5	40.7%	10.9%	13.3%	11.4%		6.1%	2.7%	0.6%	1.4%	0.8%	1.0%	3.6%	1.0%	0.9%	1.9%	0.5%	2.7%	0.4%
ROI6	30.6%	1.1%	0.3%	0.5%	3.5%		31.6%	4.0%	0.2%	1.0%	0.2%	0.6%	0.1%	1.8%	17.9%	5.9%	0.3%	0.3%
ROI7	17.9%	7.3%	0.4%	0.3%	3.0%	21.8%		15.7%	7.3%	2.3%	0.2%	0.7%	0.9%	2.5%	10.7%	6.7%	0.5%	1.8%
ROI8	31.5%	3.6%	0.5%	0.5%	1.2%	5.8%	17.6%		12.5%	6.4%	0.7%	0.9%	0.3%	0.7%	4.2%	8.6%	0.4%	4.8%
ROI9	26.7%	13.2%	2.2%	0.6%	3.1%	0.4%	17.9%	8.0%		15.0%	5.7%	0.8%	1.0%	0.5%	0.9%	0.5%	0.3%	3.2%
ROI10	28.9%	10.8%	2.0%	0.5%	0.9%	1.4%	5.7%	8.8%	15.1%		14.2%	1.3%	1.2%	0.9%	0.9%	0.8%	0.5%	6.2%
ROI11	10.8%	22.0%	4.1%	0.5%	0.6%	0.4%	0.5%	1.5%	6.1%	23.7%		14.4%	10.1%	0.7%	0.4%	1.0%	1.0%	2.0%
ROI12	26.5%	4.5%	2.0%	0.9%	6.8%	0.9%	1.1%	0.9%	0.7%	2.0%	13.5%		35.8%	1.3%	0.8%	1.0%	0.7%	0.4%
ROI13	16.0%	6.9%	18.0%	2.8%	3.0%	0.2%	0.3%	0.3%	0.3%	2.1%	9.6%	35.7%		1.7%	1.1%	0.4%	0.3%	1.2%
ROI14	42.6%	1.4%	0.9%	0.9%	3.0%	2.1%	7.1%	0.9%	0.6%	1.1%	0.6%	2.2%	1.0%		25.6%	7.8%	1.7%	0.6%
ROI15	29.2%	0.8%	1.1%	0.8%	2.7%	23.4%	12.6%	2.6%	0.3%	0.7%	0.2%	0.6%	0.7%	16.2%		6.9%	1.0%	0.2%
ROI16	25.6%	1.6%	1.1%	0.9%	1.1%	12.5%	18.7%	11.9%	0.7%	1.4%	0.6%	1.8%	0.8%	9.9%	10.2%		0.7%	0.5%
ROI17	39.6%	2.2%	5.5%	18.1%	11.1%	2.0%	1.2%	1.2%	0.8%	1.4%	0.9%	2.1%	2.8%	4.7%	4.5%	1.2%		0.6%
ROI18	21.0%	2.0%	2.5%	1.9%	1.7%	1.6%	9.8%	14.7%	16.3%	14.1%	6.5%	1.5%	0.5%	3.0%	0.9%	1.0%	0.8%	

Transition Matrix : The numbers in each row represent the relative transition probabilities from the hidden state, excluding self-transitions.

Figure 2: Gaze pattern extracted of '*Under the Wave off Kanagawa'*

indicating that points with greater depth easily lead the viewer's gaze. In the example shown in Figure 2, ROI6 just corresponds to a depth point (Mt.Fuji).

In the low rating group, no particularly distinctive gaze patterns were extracted by HMM, except for two images which have a spiral pattern.

Impression Rating; Colour

A one-sided t-test was conducted to determine whether the impression rating differed significantly between the colour and grayscale images, assuming equal variance. The results showed that the impression rating of colour images was significantly greater than that of grayscale images ($t(10)=3.15$, $p=0.002$). Furthermore, a t-test was conducted to confirm whether the means of the high and low ratings differed significantly, assuming equal variance. The results showed a significant difference in the mean of impression rating between high and low rating for both colour ($t(10)=5.76$, $p=9E-05$) and grey scale ($t(10)=4.20$, $p=0.0009$).

Although the total area of eye movement in the colour and grayscale pictures were similar, the impression rating differed significantly, indicating that colour has a stronger influence on impression rating than composition. In other words, colouration has a stronger influence on impression rating than the "visually correctness" or structural efficiency of a composition, (i.e., easily noticeable to the viewer, regardless of expertise or experience in the arts) (Locher et al., 2007).

General Discussions

The results of our experiment showed that in *ukiyo-e* landscape paintings, the area of the viewer's eye movement was larger in paintings that were rated highly as "beauty" than in those that were rated low, and certain characteristics were extracted from the trajectory of eye movement. As mentioned in the Introduction, while "beauty" is a subjective impression rating, it is not completely disparate among individuals and is considered to have a certain degree of commonality. These results shed light on the possibility that it can make viewers feel beauty through reproducible (i.e., objective) methods such as gaze guidance. Our finding is a small step toward showing that beauty has commonality and can be reproduced by an objective method (gaze guidance). In this study, qualitative analysis using eye movement data suggested that in paintings with high ratings, the gaze moves back and forth between the sign and motifs such as Mt.Fuji and signature. In addition, in the gaze patterns extracted using HMM, we confirmed spiral movements that significantly shift the gaze within four specific paintings. These results suggested the hypothesis that "paintings with high ratings have tricks that induces the viewer's gaze extensively." However, since this characteristic pattern was also observed in the two low-rated paintings, it is necessary to investigate whether and how the spiral pattern makes impact for the impressions of beauty on paintings. In addition, there could be room to examine the effect of depth as a factor in guiding eye gaze, since in all six highly rated pictures, the ROI overlapped with the point of (assumed) greatest depth. We also found that impression ratings of "beauty" and "favourability" were influenced more by colour than by composition. However, since this experiment was conducted on *ukiyo-e* paintings, which have a limited number of colours, a more detailed investigation is required.

In addition, although we used 30-second eye movement data in this experiment, we would like to explore eye movements in a more detailed time series. Our research partially achieves (while leaving some technical limitation) the bridging of different layers of data, i.e., image-processing data analysis, behavioural-level data analysis such as eye-tracking, and qualitative data such as aesthetic value judgments, and research along these lines would make new contributions to our essential understanding of aesthetics.

Acknowledgments

This study was supported by JST CREST [JPMJCR19A1]. We would like to thank Ota Memorial Museum of Art for kindly allowing us to use digital images from their collection for free.

References

Ando, H. (1859). *One Hundred Views of Mount Fuji Vol.1 (Fujimi hyakuzu 1ST edition)* (in Japanese).

Baum, L. E., Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The Annals of Mathematical Statistics. 37 (6),* 1554–1563.

Bell, D. (2004). *Ukiyo-e Explained*. Global Oriental.

Birkhoff, G. D. (1933). *Aesthetic Measure*. Harvard University Press.

Carbon, C. C. (2017). Art perception in the museum: How we spend time and space in art exhibitions. *i-Perception, 8(1),* 2041669517694184.

Christensen, B. T., Ball, L. J., & Reberc, R. (2020). Perceptual fluency effects in judgments of creativity and beauty: creative objects are perceived fluently yet they are visually complex. *Journal of cognitive psychology, 32(1),* 45-66.

Chuk, T., Crookes, K., Hayward, G, W., Chan, B, A., & Hsiao, H. J., (2017). Hidden Markov model analysis reveals the advantage of analytic eye movement patterns in face recognition across cultures. *Cognition, 169,*102-117.

Datta, R., Joshi, D., Li, J., & Wang, J.Z. (2006). Studying aesthetics in photographic images using a computational approach. *Proceedings of 9th European Conference on Computer Vision, LNCS 3953* (pp. 288-301), Springer.

Ellahi, W., Vigier, T., & Le Callet, P. (2020). HMM-based framework to measure the visual fidelity of tone mapping operators. *Proceedings of 2020 IEEE International Conference on Multimedia & Expo Workshops* (pp.1-6). IEEE.

Forrer, M. (2017). *Hiroshige*. Prestel.

Galanter, P. (2012). Computational aesthetic evaluation: Past and future. In: McCormack, J., d'Inverno, M. (Eds) *Computers and Creativity*. Springer.

Ganczarek J, Pietras K, Stolińska A, Szubielska M. (2022). Titles and semantic violations affect eye movements when viewing contemporary paintings. *Frontiers in Human Neuroscience, 16*, 808330.

Hamm, J. (1972). *Drawing scenery: landscapes and seascapes.* Perigee Book.

Harris, F. (2012). *Ukiyo-e: the art of the Japanese print*. Tuttle Publishing.

Hayn-Leichsenring, G. U., Lehmann, T., & Redies, C. (2017). Subjective ratings of beauty and aesthetics: Correlations with statistical image properties in western oil paintings. *i-Perception, 8(3),* 2041669517715474.

Kadar, E, E., & Effken, J, A. (2008). Paintings as architectural space: "Guided Tours" by Cézanne and Hokusai. *Ecological psychology, 20(4),* 299-327.

Khalighy, S., Green, G., Scheepers, C., & Whittet, C. (2015). Quantifying the qualities of aesthetics in product design using eye-tracking technology. *International Journal of Industrial Ergonomics, 49,* 31-43.

Lacoste-Badiea, S., Gagnan, A. B., & Droulers, O. (2020). Front of pack symmetry influences visual attention. *Journal of retailing and consumer services, 54,* 102000.

Lakhal, S., Darmon, A., Bouchaud, J. P., & Benzaquen, M. (2020). Beauty and structural complexity. *Physical review research, 2(2*), 022058.

Li, C., & Chen, T. (2009). Aesthetic visual quality assessment of paintings. *IEEE Journal of Selected Topics in Signal Processing, 3(2),* 236-252.

Locher, P., Krupinski, E, A., Mello-Thoms, C., & Nodine, C, F. (2007). Visual interest in pictorial art during an aesthetic experience. *Spatial vision, 21(1-2),* 55–77.

Lüttge, U., & Souza, G. M. (2019). The Golden Section and beauty in nature: The perfection of symmetry and the charm of asymmetry. *Progress in biophysics and molecular biology, 146*, 98-103.

Nagata, S. (1995). *Hokusai: Genius of the Japanese Ukiyo-e*. Translated by John Bester. Tokyo, Kodansha International.

Nojo, Y. (2015). Eye tracking based analysis of effects of motif composition on impressions of paintings: in the case of Katsushika Hokusai's "Thirty-six Views of Mount Fuji". *International Journal of Affective Engineering, 14(3),* 223-232.

Sibley, F. (1965). Aesthetic and nonaesthetic. *The philosophical review, 74(2),* 135-159.

Sibley, F. (2001). *Approach to aesthetics: collected papers on philosophical aesthetics*. In Benson, J., Redfern, B. & Cox, R, J. (Eds.), Clarendon Press.

Smith, K. J., & Smith, F, L. (2001). Spending time on art. *Empirical Studies of the Arts, 19(2),* 229-236,

Smith, L., Smith, J., Jeffrey, K., & Tinio, P. P. L. (2017). Time spent viewing art and reading labels, *Psychology of Aesthetics, Creativity, and the Arts, 11(1),* 77-85.

Spratt, E. L., Elgammal, A. (2015). Computational beauty: Aesthetic judgment at the intersection of art and science. *Proceedings of ECCV 2014 Workshops Part I, LNCS 8925* (pp. 35-53), Springer.

Tarnowski, P., Kołodziej, M., Majkowski, A., & Rak, R, J. (2020). Eye-tracking analysis for emotion recognition. *Computational Intelligence and Neuroscience*, 2020, 1-13.

Tsuji, N. (1972). Hokusai's eye and magic, *Subaru 8,* 60-87. (in Japanese)

Tsujimoto, K., & Murakami, M. (2010). The effects of change in the composition on impressions of Japanese famous picture, *ITE Technical Report*, 34(18), 35–38.

Uhlenbeck, C., & Jansen, M. (2008). *Hiroshige, Shaping the Image of Japan*. Brill Hotei.

Yang, N., Sunaga, S., Toh, K., Ihara, H. (2018). Evaluations of universal design fonts-1, *Bulletin of Japanese Society for the Science of Design, 65(1),* 51-60.

Yanulevskaya, V., Uijlings, J., Bruni, E., Sartori, A., Zamboni, E., Bacci, F., Melcher, D., & Sebe, N.(2012). In the eye of the beholder: employing statistical analysis and eye tracking for analysing abstract paintings. *Proceedings of the 20th ACM International Conference on Multimedia* (pp.349-358). ACM.