Effects of content and viewing distance on the preferred size of moving images

Masamitsu Harasawa
Science and Technology Research Laboratories, Japan Broadcasting Corporation, Tokyo, Japan
Graduate School of Information Sciences, Tohoku University, Sendai, Japan

Yasuhito Sawahata
Science and Technology Research Laboratories, Japan Broadcasting Corporation, Tokyo, Japan

Kazuteru Komine
Science and Technology Research Laboratories, Japan Broadcasting Corporation, Tokyo, Japan

Satoshi Shioiri
Research Institute of Electrical Communication, Tohoku University, Sendai, Japan

While visual size preferences regarding still objects have been investigated and linked to the "canonical size" effect—where preferred on-screen size was significantly related to objects' real-world size—the visual size preferences related to moving images of natural scenes has not been researched. In this study, we measured the preferred size of moving images of natural scenes and short duration and investigated the effect of viewing distance on size preferences. Our results showed that the preferred size varied strongly depending on content, and we found moving images' canonical size effect. The preferred size in images of *scenery* was significantly larger than in images of *persons*, and there was a positive correlation between the preferred size and the real-world physical size of the main subjects in the images. When the viewing distance was doubled, the preferred size increased about 10% as a ratio to screen size—in contrast to the findings of a previous study. While the rationale for these findings is not yet clear, our analysis suggests that neither the motion component in the images nor the nature of their background area are contributing factors. We suggest that environment, viewing distance, and screen size may contribute to this effect.

Introduction

Our behavior is often influenced by aesthetic preferences. There are numerous studies on the relationship between visual aesthetic judgments and preferences with visual features such as color (Granger, 1955; Guilford & Smith, 1959; Hurlbert & Ling, 2007; McManus, Jones, & Cottrell, 1982; Palmer & Schloss, 2010), spatial frequency (Graham & Field, 2007; Graham & Redies, 2010), orientation (Latto & Russell-Duff, 2002), and size (Konkle & Oliva, 2011; Linsen, Leyssen, Sammartino, & Palmer, 2011). Such visual preferences often affect our behavior. For example, where to sit in a movie theater, where to stand in an art gallery, or where to move to enjoy a better view of an item of interest can be related to our visual preferences about size.

Several studies have investigated visual preferences regarding object size in pictures (Bertamini, Bennett, & Bode, 2011; Konkle & Oliva, 2011; Kosslyn, 1978; Linsen et al., 2011) and demonstrated that there is a canonical size for objects, which has a certain relationship to the size of the object in the real world. Konkle and Oliva (2011) found that the preferred visual size of the picture of an object is proportional to the logarithm of its known physical size. In their experiments, participants performed several different tasks: viewing pictures of objects of different physical sizes within a frame, drawing objects, evaluating the size of imagined objects, and adjusting the size of displayed objects. All experiments consistently showed that smaller objects in the real world (e.g., strawberries or a key) were preferred to be smaller within the frame, whereas larger objects (e.g., piano or car) were preferred to be larger; this was termed *canonical size* effect. Linsen et al. (2011) showed a very similar trend in a different experiment. Their participants observed two images surrounded by square frames containing the same object in random sizes relative to the frame and chose one of the two according to their aesthetic

Citation: Harasawa, M., Sawahata, Y., Komine, K., & Shioiri, S. (2020). Effects of content and viewing distance on the preferred size of moving images. *Journal of Vision*, *20*(3):6, 1–14, https://doi.org/10.1167/jov.20.3.6.

https://doi.org/10.1167/jov.20.3.6 Received August 16, 2019; published March 24, 2020 ISSN 1534-7362 Copyright 2020 The Authors

preferences. The average size chosen for the object image was strongly correlated with the logarithm of the physical size of the object. Bertamini et al. (2011) adopted a completely different methodology and analyzed pictures of diverse sizes from two databases of artistic images of animals; they found reliable positive correlations between the physical size of the animals depicted and the sizes of their drawn or painted images.

These findings suggest that the canonical size effect is robust and does not depend on task type. Therefore, there may be an optimal size for objects to be displayed on a screen. However, previous studies used still images, and although motion is very common in our visual world and moving objects are preferred over the same standing objects (Soranzo, Petrelli, Ciolfi, & Reidy, 2018), no serious study has investigated size preference with moving images. Therefore, in this study, we investigated size preference with visual stimuli containing motion.

Meanwhile, in recent years, displays have been increasing in size and resolution. As the typical domestic environment for watching TV, including the viewing distance, has hardly changed at all, retinal images of displayed images have also been getting larger. At least some parts of the human visual system are scale variant (Campbell & Robson, 1968; Curcio, Sloan, Kalina, & Hendrickson, 1990); thus, it is possible that the impression or perceptual experience of image contents would have changed with the expansion of retinal image size. Little is known, however, about the effect of object size changes on the visual experience. According to personal communications, some TV producers reported changing their methods for shooting videos according to the expected screen size; in the case of sumo games, some typical shots of wrestlers contained their bodies above the knees when the target screen size was 40 inches (diagonal), while a similar shot included the knees when the target screen size was 85 inches. This kind of "know-how" was acquired without understanding the reasons, and there is no scientific evidence to support the use of such methods.

In the current study, we quantitatively investigated size preference for moving images using psychophysical methods to examine the effect of viewing distance. We used a visual apparatus and movie content of very high definition in an effort to remove any negative influence of coarseness in the quality of the magnified images.

Experiment 1

The preferred physical size of moving images displayed on an 85-in. liquid crystal display (LCD) was measured using psychophysical methods at two viewing distances.

Methods

Participants

Eighty-one adults (20 males and 61 females; age range 20–49 years, mean age 34.7 years, *SD* 8.2 years) participated in the experiment. They had normal or corrected-to-normal vision. Correction of vision was performed by contact lenses instead of glasses to avoid the invasion by the edges of the glasses into the visual field. All participants were paid for their participation and provided written informed consent prior to the experiment, in keeping with the 1964 Declaration of Helsinki. The study was approved by the ethics committee on human research of the Japan Broadcasting Corporation.

Apparatus

The moving images were stored in a solid state disk (SSD) recorder (HR-7512; Astrodesign, Tokyo, Japan) and displayed on an 85-in. flat panel LCD (Prototype; Sharp, Osaka, Japan) whose screen was 189 cm wide and 106 cm high and had 7,680 by 4,320 pixels of spatial resolution (dual green), 60 fps of temporal resolution, and 158 cd/m^2 of maximal luminance with a gamma value of 2.4. A computer was used to control the SSD recorder and record the responses of participants.

The participants were seated on a chair positioned in front of the display. The center point of the display and the participants' eyes were set at the same height. The viewing distance was 79 cm or 158 cm, corresponding to 0.75 H and 1.5 H, respectively, where H is a unit of length corresponding to the height of a given display monitor that is often used to describe the TV-viewing environment (International Telecommunication Union–Radiocommunication (ITU-R), 2012). At these viewing distances, the horizontal angles of view were 100 and 61 degrees, respectively. For a display with 7,680 by 3,840 pixels, 0.75 H is the distance at which two adjacent pixels subtend at an angle of 1 arc-min according to the viewer's eye. The experiment was performed separately for each participant in a dimly lit room where the frame of the display could be recognized.

Stimuli

Forty-three uncompressed moving images were used as visual stimuli. Each had a duration of 5 s and consisted of 300 frames. All the image materials were taken from movie content made for tests and demonstrations of Super Hi-Vision, which is the 8K ultra-high-definition TV (UHDTV) system (Nakasu, 2012) developed by the Japan Broadcasting Corporation. The images included sceneries, animals, and artificial objects of physical sizes ranging from

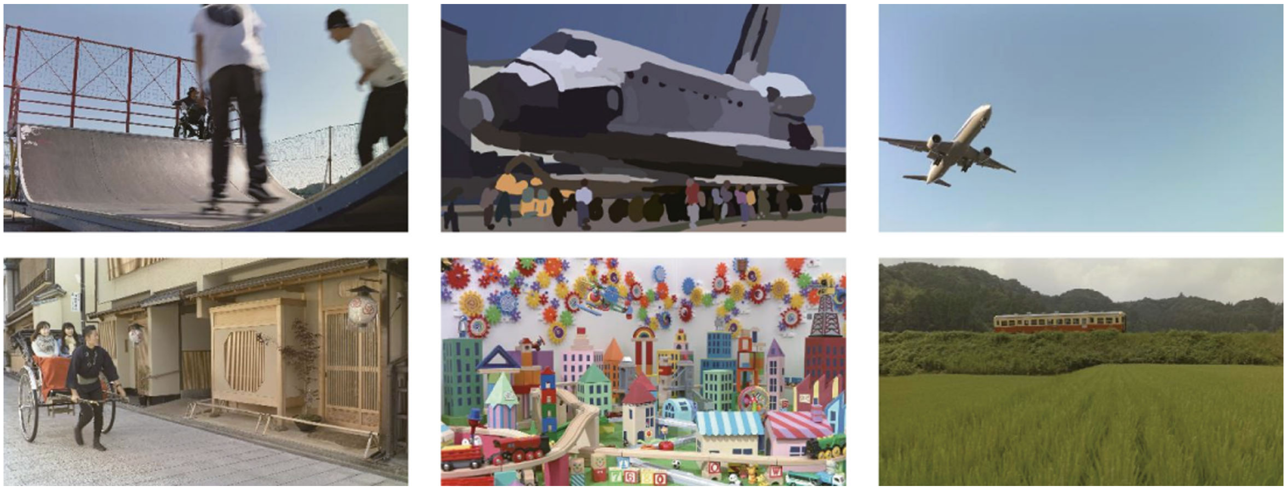

Figure 1. Examples of the moving images used as visual stimuli. All movies' aspect ratio was 16:9, and their duration was 5 s.

Figure 2. Process to produce visual stimuli. Visual stimuli were generated by shrinking the original images (7,680 by 4,320 pixels). Shrunk images were presented over a black background. Resizing factors ranged from 0.25 to 1.0 in 0.125 steps. This figure presents four out of the seven sizes.

several centimeters to a hundred meters, as well as human faces and bodies (Figure 1). In this article, the materials that were permitted to be published appear as thumbnail photos, and the other materials appear as illustrations drawn by one of the authors. Contents of the moving images appearing in this article are briefly described in the Supplementary Material.

Visual stimuli were generated from each moving image by resizing the dimensions of the original image, whose spatial resolution was 7,680 by 4,320 pixels, as follows: 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, and 1.0; thus, the smallest image was 1,920 by 1,080 pixels. The area surrounding the shrunken image was a black background (Figure 2). Each participant was presented with a total of 301 stimuli, consisting of the seven sizes of each of the 43 different moving images.

Procedures

For each moving image displayed, the participants were asked to report whether they preferred watching it in a larger or smaller physical size than that shown (i.e., two-alternative forced-choice task: shrinking or enlarging). In each trial, the visual stimulus was first displayed in the center of the display monitor for 5 s, followed by a black screen. After disappearance of the stimuli, participants reported their preference by pressing one of two buttons. The participants performed the tasks at their own pace. All stimuli were presented once in random order, and there were 301 trials in total, which were divided into four separate sessions. The participants could take several minutes to rest between sessions. Before the first session, a practice session consisting of seven or more trials was performed.

A total of 41 participants performed the experiment with a viewing distance of 0.75 H, and the remaining 40 participants did so with a viewing distance of 1.5 H. Therefore, the effect of viewing distance was examined using a between-participants design. After the psychophysical experiment, the participants provided individual information, including their age, gender, visual acuity, and TV-watching habits (size of home TV and usual viewing distance).

Analysis of data

All the data were analyzed for the two viewing distance conditions separately. A sigmoid function was fitted to the response ratios against the stimulus sizes, and the 50% response point was obtained as the preferred size for that image (Figure 3). This procedure was repeated for each moving image.

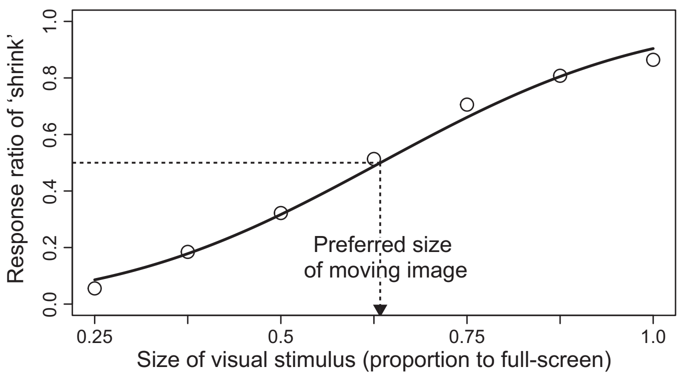

Figure 3. Calculation of preferred physical size for each movie. The 50% threshold of the ratio of "shrink" responses was defined as the preferred size for the moving image.

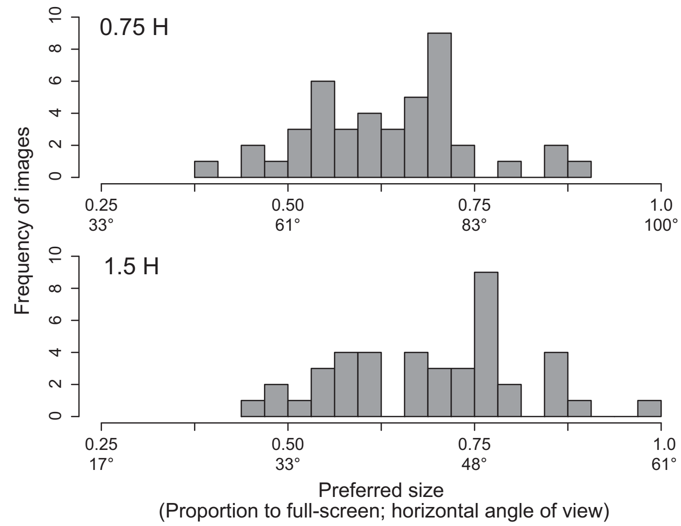

Figure 4. Histograms of preferred size of moving images in the two viewing distance conditions (upper: 0.75 H, lower: 1.5 H).

Results

Effect of content

The preferred image sizes, expressed as the ratio to the full screen, ranged from 0.40 to 0.90 of the original size in the 0.75-H viewing distance condition and from 0.46 to 1.07 in the 1.5-H condition (Figure 4). For both distance conditions, in general, the images with larger preferred size contained a scenery, or a distant object or person, and the images with smaller preferred size contained a close object or person. Figure 5 shows thumbnails of the six largest-preferred-size images and the six smallest-preferred-size images for each distance condition.

Effect of viewing distance

The distribution of preferred size in the 1.5-H condition appears displaced to the right compared to that in the 0.75-H condition (Figure 4). Statistical testing showed that this displacement was significant, $t(42) = 9.69$, $p < 0.001$, $d = 1.48$.

Figure 6 shows the effect of viewing distance on individual movies. Each small dot indicates the preferred sizes of each moving image. The preferred sizes under the two viewing distance conditions were strongly correlated, $r(41) = .95$, $p < 0.001$. The solid black oblique line indicates the line fitted by linear regression between them. Most of the dots are positioned above the dashed blue line indicating scale invariance on the screen (i.e., equal preferred size on the screen in both viewing distance conditions), which means that the preferred size on the screen increased when the viewing distance was doubled. However, this increase was much smaller than the orange dot-dash line indicating retinal

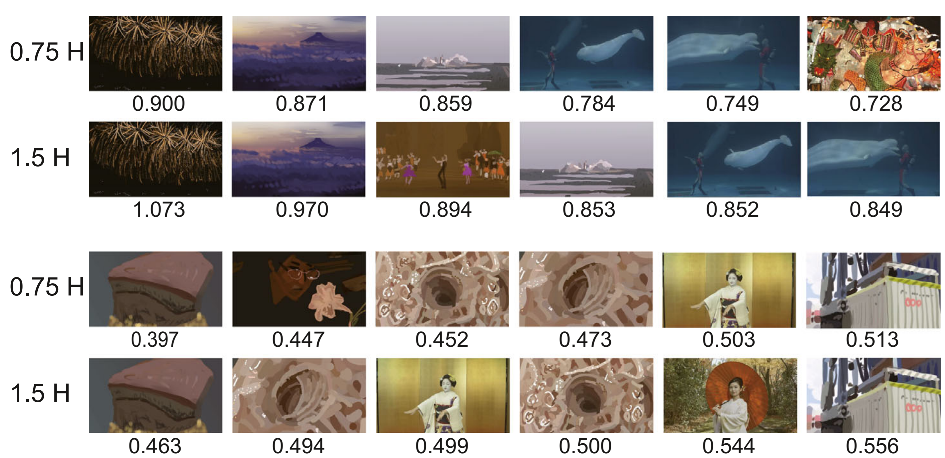

Figure 5. Images ranked according to preferred size for each viewing distance condition. The number under each image indicates the preferred size (ratio to full-screen display). For the upper 12 images, a larger size was preferred, and for the lower 12, a smaller size was selected.

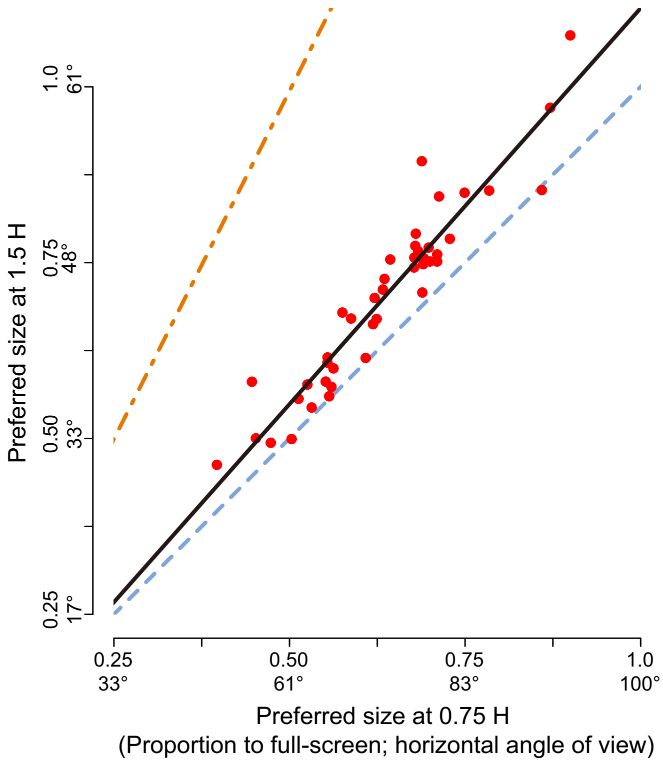

Figure 6. Preferred size of individual movies. The red dots indicate individual movies and the black slid line is the regression line. The blue dashed line indicates the scale invariance relationship on the screen (i.e., equal preferred size on the screen in both viewing distance conditions). The orange dot-dash line indicates scale variance relationship on the retina (i.e., equal preferred size on the retina in both conditions).

scale invariance (i.e., equal preferred size on the retina in both conditions).

The ratio of preferred size at 1.5 H to that at 0.75 H was distributed around 1.1 (Figure 7), with the average ratio being 1.10, meaning that at the longer distance, images 10% larger than at the shorter distance were preferred.

Individual differences

Using the same process for determining the preferred size for each moving image, the personal preferred size for each participant was determined to analyze individual differences. Since each participant observed each visual stimulus only once in order to reduce the experiment task load, it was not possible to determine the preferred size of individual moving images for individual participants. The personal preferred sizes exhibited a broad distribution (Figure 8). There was no significant difference between the two distance conditions, $t(78.98) = 1.46$, $p = 0.15$, $d = 0.32$.

These individual differences could not be accounted for by differences in participants' characteristics

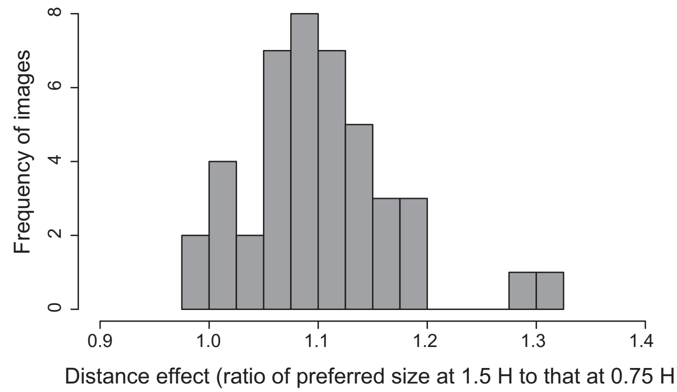

Figure 7. Histogram of distance effect (i.e., ratio of preferred size at 1.5 H to that at 0.75 H). A ratio of 1.0 means that preferred sizes were the same in both viewing distance conditions.

(Table 1): gender, age, visual acuity, size of home TVs, viewing distance from home TV, and ratio of TV size to viewing distance. After the correction of p values with the Benjamini-Hochberg method (Benjamini & Hochberg, 1995), no individual characteristic showed a significant correlation with preferred size. Therefore, there seems to be no obvious explanation for the extent of individual differences at present.

Discussion

Preferred size, which was defined as the ratio to the full screen, was obviously affected by viewing distance.

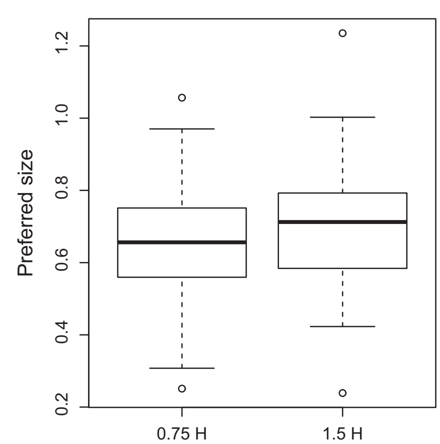

Figure 8. Distributions of preferred physical size for each participant as boxplots. The thick horizontal line indicates the median, the upper and lower edges of the box indicate top (Q1) and bottom quartiles (Q3), and the shorter horizontal lines (whiskers) above and below the box indicate the upper and lower ranges, at 1.5 times of the interquartile range to Q1 and Q3, respectively. Open circles indicate outliers, defined as data outside the range between the whiskers.

	0.75 H			1.5 H		
Variable	Individual status	p	p_{adj}	Individual status	p	p_{adj}
Gender	$t(34.3) = 0.38$, $d = 0.13$	0.709	0.773	$t(10.3) = 1.14$, $d = 0.47$	0.281	0.462
Age	$r(39) = .14$	0.383	0.511	$r(38) = -.002$	0.991	0.991
Visual acuity	$r(39) = .097$	0.544	0.653	$r(38) = -.41$	0.009	0.106
Size of home TV	$r(39) = -.16$	0.308	0.462	$r(38) = -.22$	0.167	0.462
Viewing distance from home TV	$r(39) = -.17$	0.29	0.462	$r(38) = -.34$	0.03	0.179
Ratio of TV size to viewing distance	$r(39) = .16$	0.308	0.462	$r(38) = .19$	0.24	0.462

Table 1. Results of statistical tests examining the relationship between individual status difference and the preferred size for each viewing distance condition. *Note:* p_{adj} indicates corrected p values. d indicates Cohen's d.

The average on-screen preferred size increased about 10% with a doubling of the viewing distance; this size increase was much smaller than the decrease of the retinal image.

Preferred size was also affected by the content of movies. Apparently, larger sizes were preferred for sceneries as well as distant objects and persons, whereas for magnified or close objects and persons, smaller sizes were preferred. In the following experiment, we investigated the effects of such features of movies.

Experiment 2

In order to examine the trends in preferred image size in relation to the content of movies observed in Experiment 1, participants categorized the movies according to their content, and we examined the relationship between types of content and preferred size measurements in the Experiment 1.

Methods

Participants

A total of 154 adults (76 males and 78 females; age range 20–47 years; mean age 33.5 years, *SD* 8.1 years) participated in the experiment. They had normal or corrected-to-normal vision. All the participants were paid for their participation and provided written informed consent prior to the experiment, in keeping with the 1964 Declaration of Helsinki. The study was approved by the ethics committee on human research of the Japan Broadcasting Corporation.

Apparatus

The experiment was performed in a lit room. Ten participants participated in each slot simultaneously and were seated at least 1 m from their adjacent participants. The experiments were performed with a laptop computer equipped with a 15-in. LCD. The participants watched the stimuli at the usual distance to operate laptop computers.

Stimuli

The 43 movies used in Experiment 1 were used in Experiment 2 as stimuli, as well as 72 additional movies; thus, a total of 115 movies were used in this experiment. Participants watched the movies displayed at full-screen size.

Procedures

Since the 154 participants each observed the 35 movies that were randomly selected out of the 115 movies, there was a total of 5,390 instances of exposure, and on average, each movie was shown 46.9 times.

An experimental trial started after the start button was pressed. At first, a stimulus movie was displayed, which began to play automatically and repeatedly. Participants were instructed to watch each movie at least three times for each trial and to quit playback after they had made a decision regarding the experimental question, which involved classifying the movie into one of six categories according to its content: *person*, *group of people*, *body*, *animal*, *object*, or *scenery*.

Results and discussion

In this article, only the results of the 43 movies used in Experiment 1 are presented. Figure 9 shows the response ratio of the six content categories for the movies arranged according to the mean preferred size measured in Experiment 1. There seems to be a trend that movies with smaller preferred size were often categorized as *person*, while movies with larger preferred size were often categorized as *scenery*.

Figure 10 shows the number of movies for each category (defined as the most frequent response among the six categories for each movie) and the average preferred size for each category. An analysis of variance of preferred size, except for the categories

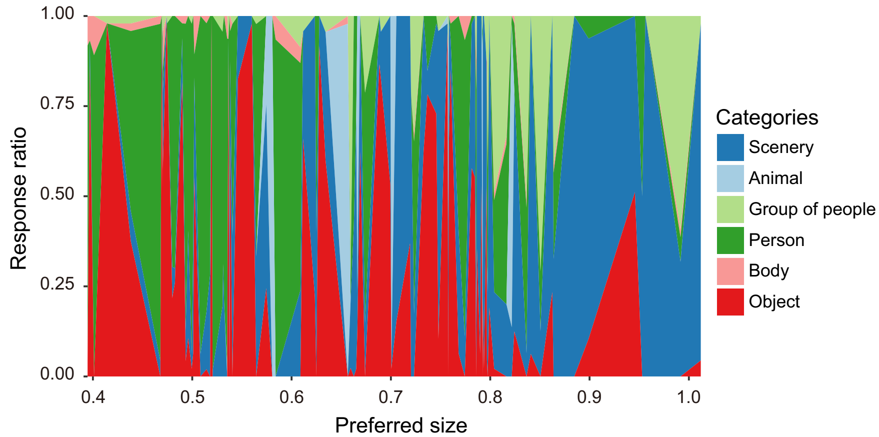

Figure 9. Results of content classification. The x-axis indicates the mean preferred size of the movies, and the y-axis represents the response ratio for each category.

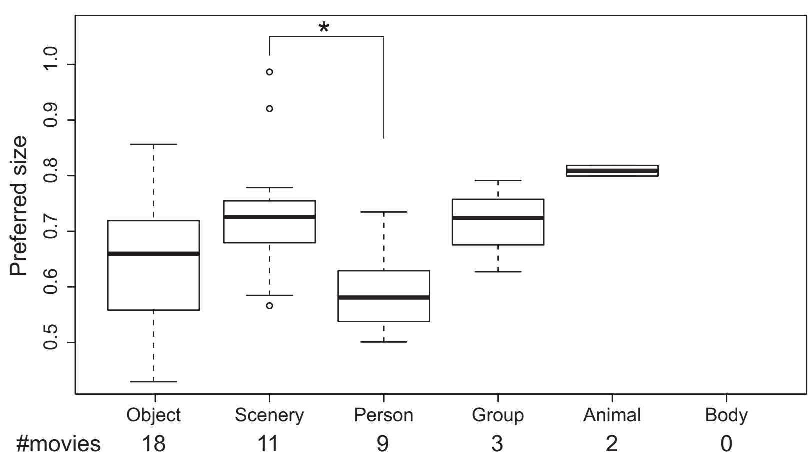

Figure 10. Preferred size for each category as a boxplot. Meanings of symbols are the same as in Figure 8. Numbers below the names of categories indicate the number of movies classified into each category. *Group of people*, *animal*, and *body* were excluded from statistical testing because of low numbers. The difference in preferred size between *scenery* and *person* was significant.

of *body*, *animal*, and *group of people*, whose number of categorized movies was zero or very few ($n < 4$), found significant variation between the categories, $F(2, 35) = 4.10$, $p = 0.025$, $\omega^2 = 0.14$. A post hoc Tukey test showed that the difference between *scenery* and *person* was significant (0.14, 95% CI [0.01, 0.26], $p = 0.028$), but the other differences were not: *person-object* (–0.04, 95% CI [–0.15, 0.07], $p = 0.65$) and *scenery-object* (0.10, 95% CI [–0.01, 0.20], $p = 0.08$). Only the ratio of *scenery* classification showed a significant correlation with mean preferred size, $r(41) = –.312$, $p_{adj} = .017$ (adjusted using the Benjamini-Hochberg method; Benjamini & Hochberg, 1995), while the other ratios showed nonsignificant correlations ($p_{adj} > .12$).

These results partially confirm the apparent trend observed in Experiment 1 that larger sizes were preferred for watching movies of sceneries and long-shots, while for movies of persons, objects, and close-shots, smaller sizes were preferred. Thus, the type of movie content seemed to affect the preferred size. However, the quantitative differences among these categories were not clearly investigated. In the following experiment, we examined the size of the main subject in each image to investigate its relationship with the preferred image size.

Experiment 3

The canonical size effect was reported to be related to the real-world size of the stimulus objects (Konkle & Oliva, 2011; Linsen et al., 2011). However, it is unclear whether the objects in our stimulus would be included in this type of preference. Therefore, we examined the relationship between the most salient object in the images and the preferred image size. For this purpose, we asked the participants to determine these features and report on their real-world sizes.

Methods

Participants

Fifty adults (25 males and 25 females; age range 23–46 years; mean age 32.8 years, *SD* 8.3 years) participated in the experiment. They had normal or corrected-to-normal vision. All the participants were paid for their participation and provided written informed consent prior to the experiment, in keeping with the 1964 Declaration of Helsinki. The study was approved by the ethics committee on human research of the Japan Broadcasting Corporation.

Apparatus

The experiment was performed in four separate sessions with 12 or 13 participants for each session. The participants used laptop computers equipped with a 15-in. LCD screen with a spatial resolution of 1,920 by 1,080 pixels. The participants observed the stimuli at the usual distance required to operate a laptop computer in a lit room.

Stimuli

Thirty movies out of the 43 movies used in Experiment 1 were used in Experiment 3 as stimuli. Participants observed the stimulus images displayed with a user interface with buttons and a slider.

Procedures

Participants were asked to determine the main subject of the motion picture by circling the area considered to be the main subject and to estimate the physical size of the area in the real world by responding with a slider. In the case of Figure 11, one participant identified the area of an aircraft as the main subject, and it can be supposed that the participant reported its size to be about 40–50 m.

A movie was played once (Figure 11A) at the beginning of a trial at full screen, followed by a responding phase (Figure 11B). In the responding phase, a still image was displayed on the top-left corner of the screen, and below the image, the participant's progress at that time was displayed. At the top-right corner of the screen, the movie was played repeatedly. Just below the movie, a progress bar was displayed together with a red upward-pointing arrow indicating the timing of the image that the participant was

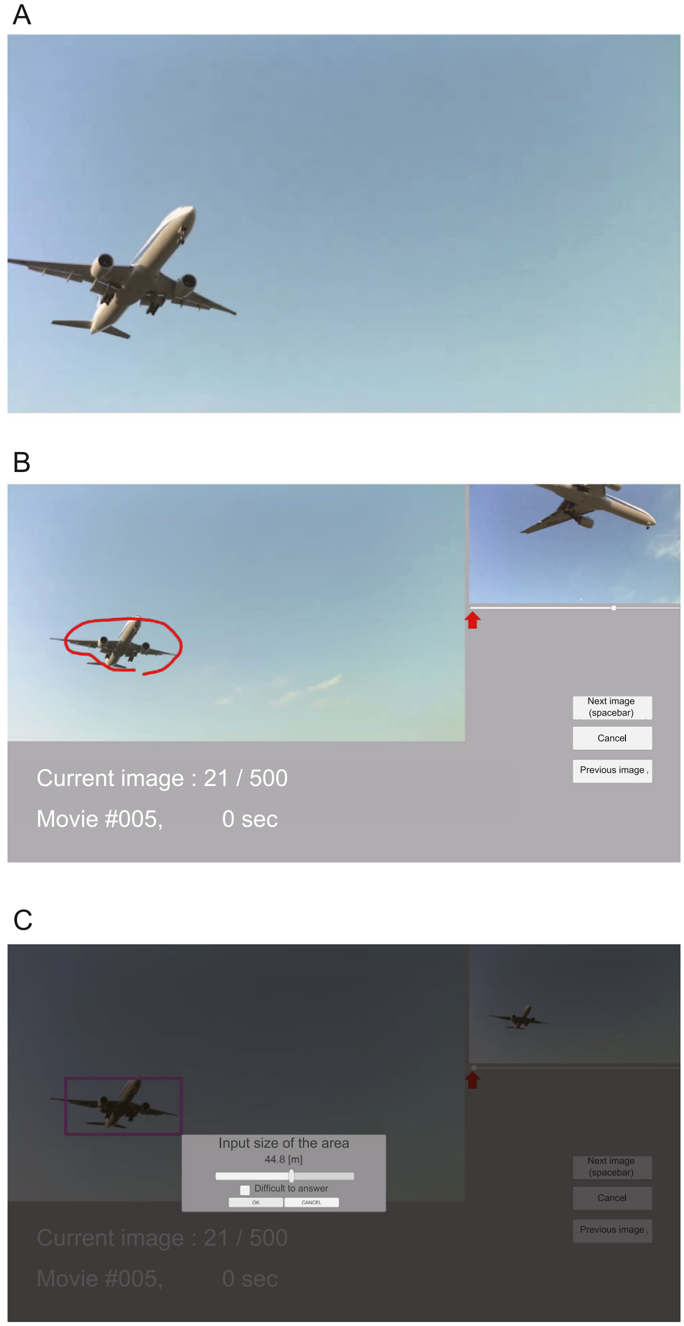

Figure 11. An example of the screens the participants observed in Experiment 2. (A) The stimulus movie was played once on the full screen at the beginning of a trial. (B) Participants drew a red circle indicating the area of the main subject of the image by dragging their mouse. (C) Participants evaluated the real-world size of the main subject and input this size by manipulating the slider. B and C were repeated five times for the image frames in intervals of 1,250 ms from the stimulus movie.

engaging with at that time. The participants were asked to circle the area considered to be the main subject of the image by dragging their mouse (Figure 11B). Immediately after the participant finished dragging the mouse, the trajectory was transformed to the smallest rectangle that could enclose the trajectory (Figure 11C). The participant was asked to report the length of the longer edge of the rectangle as a real-world physical size using a slider. The horizontal position of the sliding indicator corresponded to the logarithmic size of the area, ranging from 0.1 cm to 3.0 km, and the leftmost and the rightmost positions of the slider corresponded to "extremely small" and "extremely large," respectively. Just above the slider, a linear numeric value of the size was displayed with the unit of length. If the participants evaluated the image as having multiple main subjects, they were allowed to circle objects and report their sizes up to three times in total. After participants pressed a button labeled "Next image," the next image frame was displayed. The participants performed this process five times for different image frames whose timings were 0, 1,250, 2,500, 3,750, and 5,000 ms from the beginning.

Results and discussion

The process for calculating the estimated real-world size of the main subject for each movie includes the following steps. First, the size value for a frame was determined. Then, if only a single area was evaluated as the main subject in the frame, its estimated size was recorded. Otherwise, multiple reported size values were averaged logarithmically to show the size of representative objects. We used a logarithmic scale because previous studies suggested that the logarithmic size of the objects was involved in the canonical size effect (Konkle & Oliva, 2011; Linsen et al., 2011). Subsequently, all 250 of these values (50 participants by five frames) were averaged over participants and frames. In this averaging process, the upper and lower 10% of values were omitted to reduce the effect of outliers, and the averaging was also performed logarithmically. Consequently, the averaged real-world size of the main subject was computed for each stimulus movie.

Pearson's product-moment correlation coefficient was computed to assess the relationship between the logarithm of the real-world size of the main subject and the preferred image size as a ratio to the full screen averaged over the two viewing distance conditions. A scatterplot summarizes the results (Figure 12) and shows a strong, positive correlation between the real-world size and the preferred image size, $r(28) = 0.684$, $p = 0.00003$, which is consistent with previous studies (Konkle & Oliva, 2011; Linsen et al., 2011) that reported a correlation between the size preference and the logarithmic size of objects. The canonical size effect was replicated quantitatively for the main objects in the moving pictures.

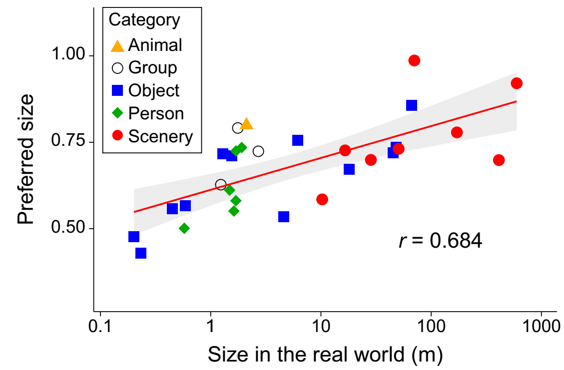

Figure 12. The relationship between the estimated size of the main subjects and the preferred image size. The x-axis indicates the estimated size in the logarithmic scale. The y-axis indicates the preferred image size measured in Experiment 1 averaged over the two viewing distance conditions. A red line represents a regression line and the gray area represents 95% confidence intervals. Symbols represent categories classified in Experiment 2 and shown in Figure 10.

In terms of the categories classified in Experiment 2, it appears that the real-world size of *Scenery* might be larger than that of *Object*. An analysis of variance of the estimated real-world size, however, showed no significant difference among the categories, $F(4, 25) = 2.01$, $p = 0.12$, $\omega^2 = 0.12$.

General discussion

We measured the preferred physical size of high-resolution moving images displayed on an 8K UHDTV using psychophysical methods. The preferred size varied with the content of movies. A change in viewing distance resulted in little or no change in the trend of variance of the preferred size related to movie contents. The average on-screen preferred size increased about 10% with a doubling of the viewing distance; this increase was much smaller than the reduction in the retinal image caused by the increase in viewing distance.

In a previous study investigating image size preference (Linsen et al., 2011), where participants observed object images of varying sizes surrounded by a rectangular frame and selected the best size according to their aesthetical preferences, the selected size was found to be proportional to the logarithm of the object's physical size. In another study, participants were required to draw a picture of a named object on a sheet of paper (Konkle & Oliva, 2011). The ratio of the size of drawn objects to the paper size was also proportional to the logarithm of the physical size of the objects. This ratio did not vary with the size of paper; therefore, the aesthetically preferred size of objects seemed to be linked to the size of the surrounding frame. In a different context, Hubbard and Baird (1988)

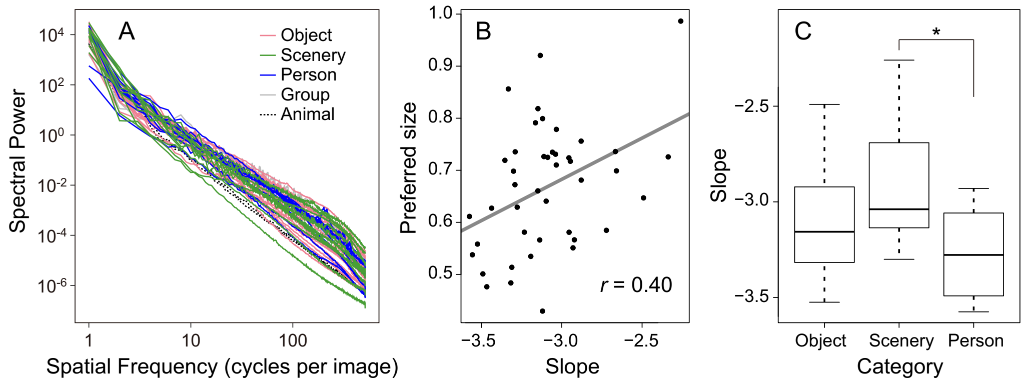

Figure 13. (A) Fourier power spectra. In the log-log plane, Fourier power was plotted as a function of spectral frequency. The power spectra were computed by radially averaging the two-dimensional Fourier power. Curves indicate the spectrum of each movie, and they are colored according to the categories determined in Experiment 2. (B) The relationship between the preferred image size and the slope of the regression line of the power spectrum. The dots represent each movie. (C) Distributions of the slopes of the regression lines for the three categories.

quantitatively investigated the relationship between the physical size of objects and the distance at which they were spontaneously imagined ("first-sight" distance) and found evidence for a power-law relationship between object size and first-sight distance. Since viewing distance is inversely related to the retinal object size, this finding would be consistent with the two abovementioned studies (Konkle & Oliva, 2011; Linsen et al., 2011).

In our current study, larger preferred sizes were often observed for sceneries and distant objects and persons, while smaller sizes were often preferred for magnified or close objects and persons. This trend seems consistent with previous studies (Konkle & Oliva, 2011; Linsen et al., 2011) and was confirmed quantitatively, at least partially, by the significant difference in preferred size between the two types of stimuli, *scenery* and *person*. Since sceneries generally contain very large objects (e.g., mountain, sea, and rice field) and are obviously larger than persons, our data seem consistent with previous findings.

This observation was for whole images, not for individual objects in scenes; we examined the relationship between the size of the main objects of the images and the size preference and found a significant positive correlation, the canonical size effect (Konkle & Oliva, 2011; Linsen et al., 2011) in moving pictures.

We also examined the effects of three image statistics on the preferred image size: average luminance, deviation of luminance variable, and spatial frequency distribution. For calculation of the luminance image statistics, images were resized to 1,920 by 1,080 pixels from the original size, and statistics were averaged over 10 image frames extracted in intervals of 0.5 s for 300 frames. Not surprisingly, Pearson's product-moment correlation coefficients with the preferred image size were not significant for mean luminance, $r(41) = -0.12$, $p = 0.43$, and standard deviation of luminance, $r(41) = -0.06$, $p = 0.73$.

Spatial frequency distribution was characterized by the slope of the power spectrum because the logarithmic value of Fourier power spectra of natural images tends to be inversely proportional to the power of spatial frequency as f^{-s}, where f and s indicate spatial frequency and a constant (Burton & Moorhead, 1987; Field, 1987; Tolhurst, Tadmor, & Chao, 1992; van der Schaaf & van Hateren, 1996). We calculated the Fourier power spectra as radially averaged two-dimensional Fourier power computed from the 1,024-by-1,024-pixel images resized from the 3,840-by-3,840-pixel center square area of the original images. We averaged the power spectra over 10 image frames extracted in intervals of 0.5 s for each motion picture, similar to the processes for luminance statistics (Figure 13A). Subsequently, the slopes of the regression lines for the spectral power against the spatial frequency in the log-log domain were computed. Pearson's product-moment correlation coefficient between the slopes and the preferred image size was significant, $r(41) = 0.40$, $p = 0.008$ (Figure 13B). In addition, the slopes showed a significant difference among the categories determined in Experiment 2 based on an analysis of variance with those three categories, $F(2, 35) = 3.68$, $p = 0.035$, $\omega^2 = 0.12$, and a post hoc Tukey test showed that the difference between *scenery* and *person* was significant (0.35, 95% CI [0.03, 0.67], $p = 0.032$), but the other differences, *person-object* (–0.13, 95% CI [–0.43, 0.15], $p = 0.50$) and *scenery-object* (0.21, 95% CI [–0.06, 0.49], $p = 0.15$), were not (Figure 13C). This trend regarding the image statistics is consistent with that found in a previous study (Redies, Hänisch, Blickhan, & Denzler, 2007) that reported that

Figure 14. Distance from the regression line and ΔL. The upper row of numbers below images are the distances from the regression line of preferred sizes between the two viewing distance conditions in Figure 6. Positive numbers indicate results above the line and negative numbers below the line. The lower row of numbers below the images indicate ΔL, a motion-related index calculated as the root mean square of the pixelwise difference in luminance between adjacent motion frames. The upper and lower rows of images have the highest and the lowest values of distance, respectively.

photos of human faces showed a steeper slope than natural scenes did. Therefore, it is plausible that the relationship between the preferred image size and the categories found in Experiment 2 could be accounted for by the spatial frequency characteristics, at least partially. However, since the correlation coefficients of the preferred size to the power spectrum distribution were smaller than those to the evaluated size of the main objects, $r_{\text{slope}} = 0.39$, $r_{\text{size}} = 0.68$, $z = 1.69$, $p = 0.090$ (Pearson & Filon, 1898), it does not seem that spatial frequency distribution plays a deterministic role exclusively, and it would not be reasonable to rule out the idea that the size of main objects in the images played some role related to the participants' size preference. This view was supported by the result of the multiple regression analysis testing if the spatial frequency distribution (the slopes of the regression lines of the power spectra) and the evaluated size of the main subject predicted the preferred image size. The results of the regression indicated the two predictors explained 49.1% of the variance [adjusted $R^2 = 0.491$, $F(2, 27) = 15$, $p < 0.0001$] and found that the evaluated size significantly predicted the preferred size ($\beta = 0.085$, 95% CI [0.047, 0.122], $p < 0.0001$) but the slope did not ($\beta = 0.112$, 95% CI [–0.014, 0.239], $p = 0.08$).

The difference in preferred size involved in changing the viewing distance was much smaller than that expected in terms of scale invariance on the retina (Figure 6). This result could suggest that preferred size would be basically linked to the physical size of the frame rather than the retinal size. This trend also seems consistent with a previous study (Konkle & Oliva, 2011). The similarities between the current and previous studies suggests that our work expands the findings regarding still single objects by also including results of natural scenes and moving images.

The effect of viewing distance was modest but not negligible. Our data showed that the preferred size on the screen increased about 10% with a doubling of the viewing distance, which reduced the size of the frame on the retina by a half. Konkle and Oliva (2011), however, showed that the change of paper size that induced the change of the frame size on the retina did not change the size ratio of drawn objects to paper. Thus, changing the frame size on the retina affected preferred size in our data but did not in Konkle and Oliva (2011). There are two plausible accounts for this discrepancy. One involves the method of changing the frame size. While Konkle and Oliva (2011) changed the size of paper but did not change the viewing distance, we changed the viewing distance but did not change the size of the screen. This suggests that the frame of the screen would not always play a role as a reference frame for comfortable size in visual object representation. If comfortable size could be determined as a ratio to some virtual reference frame, it is supposed that the reference frame could not be uniquely determined either on the retina or on the screen. The other factor involved in the frame size effect would be the content of stimuli. While previous research used still single objects as stimuli, we used movies of natural scenes. Therefore, our stimuli included a background area surrounding the main subject as well as motion components.

In order to investigate the effect of the background area surrounding the main subject, we compared the images that showed the largest and the smallest viewing distance effects. Figure 14 shows the images and their distance from the regression line, which is shown in Figure 6 as a black solid line. Larger distance values indicate a larger viewing distance effect; preferred size was much larger in the longer than in the shorter viewing distance condition. In the upper and lower rows in Figure 14 are the five images with the highest and lowest distance values, respectively. An obvious difference between the images in the upper and lower rows was not found regarding background areas surrounding the object assumed to be the main subject in each image. Some images have many objects in the background area, while other images have sparse background areas. Thus, the background area did not seem to have been involved in the different trends in terms of the frame size effect.

A previous study investigating the effect of screen size and viewing distance on visually induced motion sickness (Shigemasu et al., 2006) revealed that some components of the motion sickness symptoms were affected by the screen size, while others were affected by the retinal image size. It seems possible that increased feelings of discomfort related to visually induced motion sickness could be due to an increase in motion intensity caused by the enlarged retinal image at a shorter viewing distance, and such feelings could be involved in the preference for smaller size at shorter viewing distance.

In order to investigate the effect of the motion component contained in our visual stimuli on preferred size, we examined the relationship between several statistics derived from preferred size and ΔL, an index of motion, calculated as the root mean square of pixelwise difference of luminance between adjacent motion frames by

$$\Delta L = \sqrt{\frac{1}{(f-1)\cdot w \cdot h}\sum_{t=2}^{f}\sum_{x=1}^{w}\sum_{y=1}^{h}(L_{x,y,t} - L_{x,y,t-1})^2} \quad (1)$$

where $L_{x,y,t}$ is the luminance at position x, y in the frame t, w and h are the width and height of the image, respectively, and f is the number of frames, 300. ΔL was obtained for each movie resized to 1,920 by 1,080 pixels.

We examined the correlation of ΔL or logarithm of ΔL with several statistics derived from preferred size (i.e., average over viewing distances, ratio, and difference between the two viewing distances). No significant correlations, however, were observed ($ps > .27$). Figure 14 shows the images with the highest and lowest values of distance from the regression line, indicating the viewing distance effect. No obvious trend was found in the relationship between the viewing distance effect and the motion component represented by ΔL. Thus, at least at the moment, we could not identify an effect of motion contained in movie stimuli.

The change in preferred size with viewing distance could be considered in some respects the result of failure in size constancy (Koh & Charman, 1999; Lawson, Bertamini, & Liu, 2007; Norman, Todd, Perotti, & Tittle, 1996). If size constancy worked perfectly, preferred size would not vary with viewing distance and would be constant on the screen, and if size constancy did not work at all, preferred size would be constant on the retina. Our results actually fall between these two scenarios.

According to the above discussions, the change in preferred size of about 10% induced by the viewing distance variation may be due to a difference in estimated distance. The detailed rationale, however, has not been clarified so far, and further investigation is needed.

A previous study using still images (Linsen et al., 2011) reported that the relationship between preferred size and real-world size was weakened by a reduction in the amount of image detail. It seems plausible that a high-resolution display, such as the 8K UHDTV, may underlie the similar size relationship observed in our study. From the perspective of application, since visual displays have been becoming larger and denser in recent years, such a size effect may be more salient than before and become an important issue in movie production.

Conclusions

We measured the preferred physical size of short-duration moving images presented on a large high-resolution display at two different viewing distances, namely 0.75 and 1.5 times the height of the monitor. The results showed that the preferred size of the images varied strongly depending on their content; in particular, the real-world physical size of the main object contained in the images seemed to be involved in the participants' preferences. The preferred size at the longer viewing distance was 10% larger than at the shorter viewing distance. This difference was not observed in previous research on size preference using still images of single objects. The detailed rationale of such a trend was not elucidated; however, the presumed cause would not be related to the content of images, motion component, or background area but to the environment, viewing distance, and screen size.

Keywords: visual aesthetic preference, size, moving image, natural scene, viewing distance, psychophysics

Acknowledgments

We thank Editage (www.editage.com) for English language editing.

Commercial relationships: none.
Corresponding author: Masamitsu Harasawa.
Email: harasawa.m-ii@nhk.or.jp.
Address: Science and Technology Research Laboratories, Japan Broadcasting Corporation, Tokyo, Japan, and Graduate School of Information Sciences, Tohoku University, Sendai, Japan.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*, 289–300.

Bertamini, M., Bennett, K. M., & Bode, C. (2011). The anterior bias in visual art: The case of images of animals. *Laterality: Asymmetries of Body, Brain and Cognition*, *16*, 673–689, https://doi.org/10.1080/1357650X.2010.508219.

Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, *26*, 157, https://doi.org/10.1364/AO.26.000157.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*, *197*, 551–566.

Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *The Journal of Comparative Neurology*, *292*, 497–523, https://doi.org/10.1002/cne.902920402.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America*, *4*, 2379, https://doi.org/10.1364/JOSAA.4.002379.

Graham, D. J., & Field, D. J. (2007). Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision*, *21*, 149–164, https://doi.org/10.1163/156856807782753877.

Graham, D. J., & Redies, C. (2010). Statistical regularities in art: Relations with visual coding and perception. *Vision Research*, *50*, 1503–1509, https://doi.org/10.1016/j.visres.2010.05.002.

Granger, G. W. (1955). An experimental study of colour preferences. *The Journal of General Psychology*, *52*, 3–20, https://doi.org/10.1080/00221309.1955.9918340.

Guilford, J. P., & Smith, P. C. (1959). A system of color-preferences. *The American Journal of Psychology*, *72*, 487–502, https://doi.org/10.2307/1419491.

Hubbard, T. L., & Baird, J. C. (1988). Overflow, first-sight, and vanishing point distances in visual imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 641–649, https://doi.org/10.1037/0278-7393.14.4.641.

Hurlbert, A. C., & Ling, Y. (2007). Biological components of sex differences in color preference. *Current Biology*, *17*, R623–R625, https://doi.org/10.1016/j.cub.2007.06.022.

International Telecommunication Union–Radiocommunication (ITU-R). (2012). *General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays (Recommendation No. BT.2022)*. Retrieved from https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2022-0-201208-I!!PDF-E.pdf

Koh, L. H., & Charman, W. N. (1999). Size constancy and angular size matching in size perception of near objects. *Optometry and Vision Science: Official Publication of the American Academy of Optometry*, *76*, 241–246.

Konkle, T., & Oliva, A. (2011). Canonical visual size for real-world objects. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 23–37, https://doi.org/10.1037/a0020413.

Kosslyn, S. M. (1978). Measuring the visual angle of the mind's eye. *Cognitive Psychology*, *10*, 356–389, https://doi.org/10.1016/0010-0285(78)90004-X.

Latto, R., & Russell-Duff, K. (2002). An oblique effect in the selection of line orientation by twentieth century painters. *Empirical Studies of the Arts*, *20*, 49–60, https://doi.org/10.2190/3VEY-RC3B-9GM7-KGDY.

Lawson, R., Bertamini, M., & Liu, D. (2007). Overestimation of the projected size of objects on the surface of mirrors and windows. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1027–1044, https://doi.org/10.1037/0096-1523.33.5.1027.

Linsen, S., Leyssen, M. H. R., Sammartino, J., & Palmer, S. E. (2011). Aesthetic preferences in the size of images of real-world objects. *Perception*, *40*, 291–298, https://doi.org/10.1068/p6835.

McManus, I. C., Jones, A. L., & Cottrell, J. (1982). The aesthetics of colour. *Perception*, *10*, 651–666.

Nakasu, E. (2012). Super hi-vision on the horizon: A future TV system that conveys an enhanced sense of reality and presence. *IEEE Consumer Electronics Magazine*, *1*, 36–42, https://doi.org/10.1109/MCE.2011.2179821.

Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S. (1996). The visual perception of three-dimensional length. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 173–186.

Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 8877–8882, https://doi.org/10.1073/pnas.0906172107.

Pearson, K., & Filon, L. N. G. (1898). Mathematical contributions to the theory of evolution: IV. On the probable errors of frequency constants and on the influence of random selection on variation and correlation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *191*, 229–311, https://doi.org/10.1098/rsta.1898.0007.

Redies, C., Hänisch, J., Blickhan, M., & Denzler, J. (2007). Artists portray human faces with the Fourier statistics of complex natural scenes. *Network: Computation in Neural Systems*, *18*, 235–248, https://doi.org/10.1080/09548980701574496.

Shigemasu, H., Morita, T., Matsuzaki, N., Sato, T., Harasawa, M., & Aizawa, K. (2006). Effects of physical display size and amplitude of oscillation on visually induced motion sickness. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology - VRST '06*, 372–375, https://doi.org/10.1145/1180495.1180571.

Soranzo, A., Petrelli, D., Ciolfi, L., & Reidy, J. (2018). On the perceptual aesthetics of interactive objects. *Quarterly Journal of Experimental Psychology*, *71*, 2586–2602, https://doi.org/10.1177/1747021817749228.

Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, *12*, 229–232, https://doi.org/10.1111/j.1475-1313.1992.tb00296.x.

van der Schaaf, A., & van Hateren, J. H. (1996). Modelling the power spectra of natural images: Statistics and information. *Vision Research*, *36*, 2759–2770, https://doi.org/10.1016/0042-6989(96)00002-8.