Journal of Imaging Science and Technology® 61(6): 060401-1–060401-12, 2017.

Evaluation for Faithful Reproduction of Star Fields in a Planetarium

Midori Tanaka▲
College of Liberal Arts and Sciences, Chiba University, Chiba, Japan
E-mail: midori@chiba-u.jp

Takahiko Horiuchi▲
Graduate School of Engineering, Chiba University, Chiba, Japan

Ken'ichi Otani
Engineering Division, Konica Minolta Planetarium Co., Ltd., Tokyo, Japan

Po-Chieh Hung▲
Ex-Konica Minolta, Inc., Tokyo, Japan

Abstract. *In order to investigate factors necessary for reproducing actual star images in a planetarium, for this article, the authors conducted a psychophysical experiment using projection stimuli generated by changing three parameters of the stars: color, luminance, and size. A reference projection pattern was designed to be faithful to the actual starry sky perceptually (rather than physically) by an experienced group with abundant astronomical observation experience. A reproduction system was constructed to project ten types of star image patterns to a planetarium dome using different parameters. Then, evaluation experiments with twenty observers were conducted. The results of the experiment indicate that the intensity of the stars was sensitive to the fidelity of the reproduction, and in either case of change (whether the star was bright or dark compared to the reference pattern), the result was a loss of fidelity. In addition, although the fidelity was improved when the size of the projected star was small, for stars that were projected larger than the reference pattern, the result was remarkably negative. As for differences in color, the evaluation results suggested that the tolerance to loss of fidelity was wide.*

[DOI: 10.2352/J.ImagingSci.Technol.2017.61.6.060401]

INTRODUCTION

Illumination is essential to enrich the lives of people, and provides us with a safe, comfortable living environment even on dark evenings. On the other hand, places where we can observe a skyful of stars have become precious and rare. This is because the excessively rich illumination in our environment has made it impossible to witness the beauty of thousands of sparkling stars in the night sky due to light pollution.[1,2] For this reason, together with the development of modern illumination, has been the development of planetariums to reproduce a starry sky artificially using an image reproduction system. These serve those in fields such as astronomy education and entertainment, and have played an important role in communicating the majesty of universal (deep) space.[3] Many studies have been reported on methods of observation and acquisition of images of the starlit sky,[4] and on using computer graphics (CG) reproduction methods in the display.[5,6] However, there has been insufficient discussion about methods appropriate for reproducing star images and assessing such methods in a planetarium that imitates starry sky.

In order to give an impression equivalent to a real starry sky with a projected image in a planetarium, it is necessary to consider the viewing environment and the visual characteristics involved in astronomical observation. Astronomical observations in the real world, and observations of star images in a planetarium, are mostly observed using scotopic vision; however, it is thought that we use partially photopic or mesopic vision systems also because we can perceive actual star colors. Mesopic vision is active during the process of transition from photopic (illuminated) to scotopic (dark) vision, and utilizes a complicated mechanism for perceiving color because both cones and rods work together. For example, the peak of human spectral sensitivity to perception of brightness shifts to shorter wavelengths with declining intensity.[7] Mesopic vision has special characteristics, such as the Purkinje shift by which red colors (long wavelength) are perceived as darker than blue colors (short wavelength). In recent years, there have been studies on methods to reproduce images in the mesopic vision environment that consider optic nerve function.[8–10] There have also been psychophysical experiments for distinguishing colors in dim light (low illumination) environments.[11,12] For example, Shin et al.[12] reported that reduction of brightness level caused not only hypofunction of saturation and brightness perception, but also shift of hue perception under different illumination environments eliciting photopic, mesopic, and

---

▲ IS&T Members.

Received June 21, 2017; accepted for publication Sept. 18, 2017; published online Nov. 28, 2017. Associate Editor: Rita Hofmann-Sievert.

1062-3701/2017/61(6)/060401/12/$25.00

scotopic vision, when color patches were displayed on a CRT monitor.

Incidentally, star images in a planetarium (which are the subject of this study) are a set of spatially distributed point-like light sources having various color, luminance, and size. Moreover, the viewing environment is unique and different from other image reproduction situations, because a field of stars must be projected onto a hemispheric screen representing the entire sky, which is different to general display devices such as plane screens with a limited field of view. In addition, although a starry sky is three-dimensional, it is considered two-dimensional perception, approximately the same as a normal image due to the extremely long viewing distance between the stars and observers. For the above-mentioned reasons, many of the findings from conventional experimental studies on display of natural images and patches may not necessarily be applicable to reproduction of star images in a planetarium. Furthermore, the reproduction of star images in a planetarium has limited reproduction performance for color, luminance, contrast, resolution, and dome size in comparison with general image output devices such as monitors. Thus, it does not seem to be easy to provide an image equivalent to an actual starry sky.

For the work reported in this article, we investigated factors that indicate the faithfulness of starry sky reproduction in a planetarium with focus on star color, luminance and size. This was done by conducting a psychophysical experiment with human observers, in which important indices used to represent the faithful reproduction of stars were analyzed.

EXPERIMENT

In our experiment, we analyzed the factors that influence faithful reproduction of a starry sky by evaluating the faithfulness of the images reproduced in planetarium projections by changing each of three parameters (color, luminance, and size) of individual stars.

Reproduction method of stars

The projection systems that produce star image in a planetarium are mainly either optical or digital systems. In our preliminary experiment to compare the image reproduction of both systems, it was difficult to reproduce enough resolution to evaluate the image faithfulness using a digital system. Therefore, in our experiment, we used a planetarium incorporating an optical system with star plates to reproduce fixed stars, as shown in Figure 1. In addition, we prepared another projection device to produce bright stars with particular brightness magnitude to ensure enough dynamic range of luminance. High-powered, white LEDs were used to ensure adequate intensity of fixed and bright stars, as much as possible. Stars were projected on the dome screen by passing light through star plates installed in a lens barrel. The color and luminance of the stars were adjusted by inserting transmission filters in front of the light sources, and the star size was controlled by the hole diameter size of star plates.

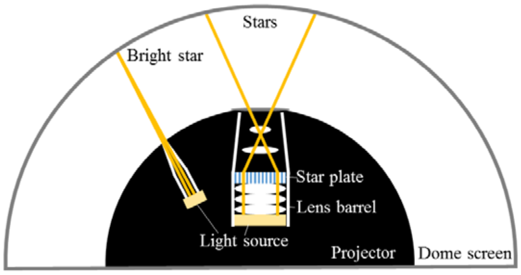


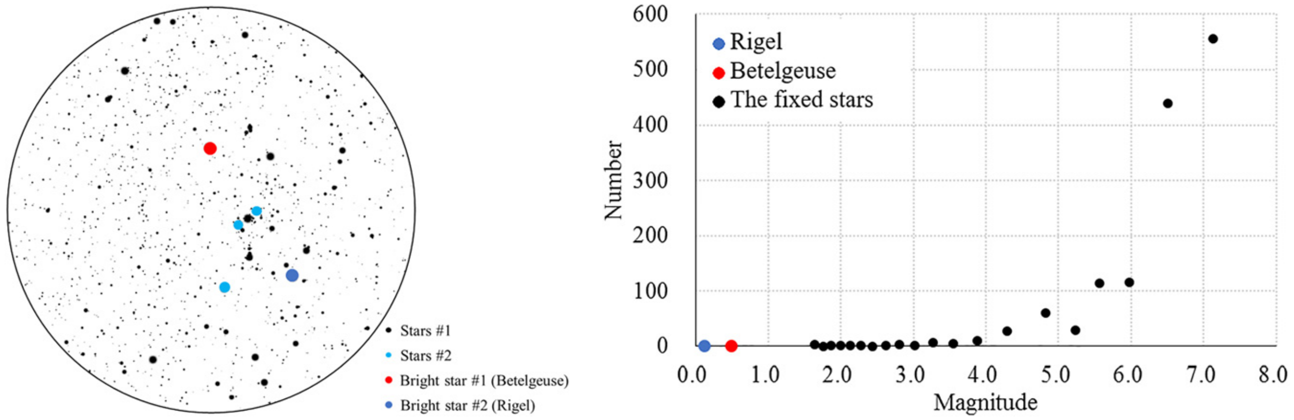
Figure 1. The construction of the optical planetarium used in our experiment.

Experimental stimuli

We selected stars around Orion (1/32 of the entire night sky) as experimental stimuli to evaluate the faithfulness of the reproduced images. Orion is a familiar constellation among Japanese people and is mainly observed during winter in Japan. Sirius, the Milky Way, and some nebulas, were originally included in the projection area. However, we excluded Sirius as a projection target because it is the brightest star and may have produced bias during evaluations due to its attractiveness. In addition, we also excluded the Milky Way and nebulas due to the difficulty of managing the hole diameter for them.

For faithful reproduction of a star image, we should originally consider the physical factor of the observation environment like the twinkling of stars caused by atmospheric extinction.[13] However, we excluded external factors such as the atmospheric extinction in this experiment and determined that it was better that a stable star image was displayed temporally. This was because the twinkling could produce too much complexity and uncertainty for the observation due to its attractiveness. For similar reasons, to prevent the influence on star perception of airlight from the solar system (zodiacal light) and airlight from Earth's atmosphere (light pollution), the experiments were conducted in complete darkness to reproduce environment of scotopic vision, which is typical of a general planetarium environment. As described in Introduction, however, due to the brightness of the cluster of stars, it will be perceived partially in photopic or mesopic vision systems. We show the projected star image used as the experimental stimulus in Figure 2(a). The color of stars around Orion were represented using four different color reproduction methods, which consisted of two fixed star colors (Stars #1 and #2) and two bright star colors (Bright star #1 and #2). These were reproduced by controlling the magnitude and color temperature of each star. Bright star #1 was for Betelgeuse and #2 was for Rigel. The light source of a bright star could be projected with brightness approximately 15 times that of the light source of fixed stars.

It has been reported[14] that people can recognize stars of magnitude 6.0 with naked eyes. However, people can perceive the brightness of darker stars with magnitude more than 6.0, although eyes cannot resolve the dim star as a point. Based on this knowledge, for our experimental stimulus, we used a set of 1378 fixed stars and two additional bright stars (Betelgeuse



(a)

(b)

Figure 2. The projected experimental stimulus. (a) Representation of the color variation of stars used in our experiment. (b) The number of reproduced stars.

and Rigel) having magnitude less than 7.4. This was because the hole diameter size needed to represent magnitude 7.4 stars was at the limit of the hole processing for the star plate. Fig. 2(b) shows the number of stars corresponding to the brightness of each magnitude.

We prepared ten experimental patterns in total by changing the color and luminance of four star representations relative to the whole star image. The change of star color and luminance was implemented using ND filters and color-temperature-change filters on the lens barrel of the projector. Table I shows the list of projection patterns used as experimental stimuli. As written previously, we assumed that the color, luminance, and size of a star influenced the evaluation of the faithfulness of the star image reproduced in the planetarium. Therefore, we first determined a standard pattern (Std) that had perceptual faithfulness of an actual starry sky. Second, for the standard pattern, we prepared additional patterns in which an individual parameter had been changed. Specifically, there were three pattern (C-1–C-3) for shifted color temperature, three pattern (L-1–L-3) for changed luminance, and two pattern (S-1–S-2) for changed projection size. We added one additional experimental pattern (B) that approached the chromaticity of the Planckian locus of the real star color. We give a detailed explanation of each pattern in the following subsections.

Standard pattern

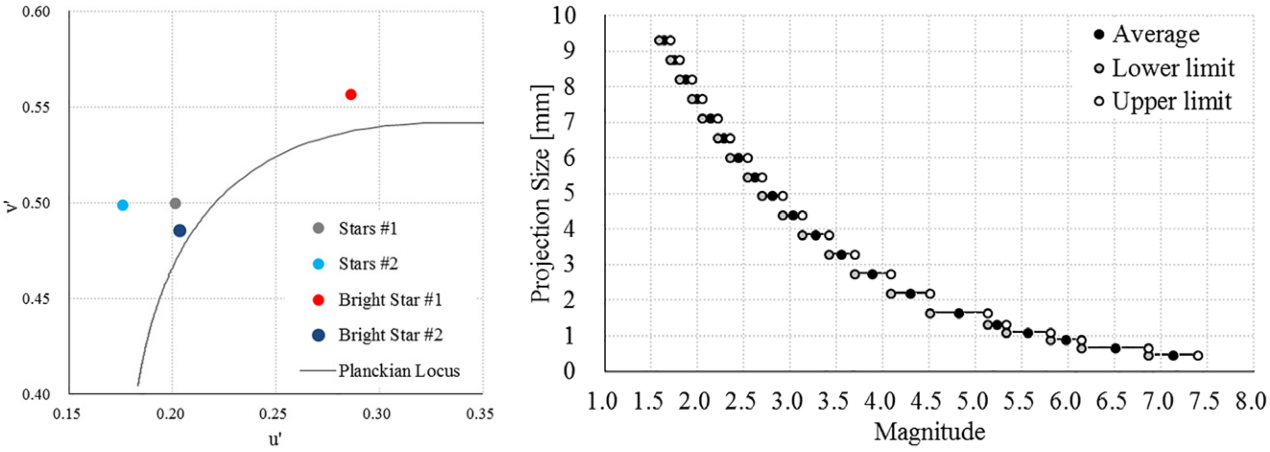
For reproduction of a real starry sky faithful enough to give the same impression as astronomical observation in real life, it is ideal to reproduce equivalent physical factors such as star color, tone, size, and depth among stars. However, reproduction of star images that are equal to those of a real sky is very difficult physically. This is because the planetarium has limited resources including light source, construction design, dome size, and optical performance of the projector. Therefore, we determined what projection conditions could reproduce star images in a planetarium that were perceptually the same as a real starry sky.

Table I. List of experimental projection patterns.

Projection Pattern	Changed Parameter	Remarks
Std	Standard	
C-1	Color-temperature shift	Pattern Std -100 mired
C-2	Color-temperature shift	Pattern Std +100 mired
C-3	Color-temperature shift	Pattern Std +200 mired
L-1	Luminance shift	1/2× Pattern Std
L-2	Luminance shift	2× Pattern Std
L-3	Luminance shift	3× Pattern Std
S-1	Size	2/3 of Pattern Std
S-2	Size	3/2 of Pattern Std
B	Color	Planckian locus

The reproduction of the color and luminance of individual stars in the standard pattern was designed to provide equivalent perception of major stars obtained from an actual starry sky by a star image projected by the planetarium, as determined by five experienced observers. They were males averaging 50 years with abundant experience of astronomical observation. The observers memorized the luminance, color, and size of actual stars after astronomical observation from a Japanese mountain under the condition of non-light polluted and clear sky. Then the standard experimental pattern was determined by memory matching with various star images in which the luminance, color, and size had been changed. This was done to overcome the difficulty of side-by-side comparison of reproduced images with the actual starry sky because the planetarium projection system is not very mobile. Figure 3(a) shows the chromaticity of each star in the standard pattern.

The reproduction of the luminance of other stars was designed in consideration of Pogson's equation which defines the magnitude now used to express the brightness of astronomical objects as a logarithmic scale.[15] The magnitude



(a) (b)

Figure 3. Condition of star reproduction in the standard pattern. (a) Star color. (b) Projection size for each magnitude of stars.

is defined as in Eq. (1).

$$m_2 - m_1 = -2.5\log_{10}(b_2/b_1). \tag{1}$$

Here, m and b mean the magnitude and brightness, respectively.

In our work, the brightness for each star magnitude was controlled perceptually by changing the hole diameter size of the star plates because it was not easy to control the brightness of the stars individually using a light source for fixed stars. The size of an actual star is too small with the infinitesimal of viewing angle. A hole diameter size capable of generating stars of equivalent size is desired for faithful representation. However, perception becomes more difficult because a star image with a small hole diameter size is too dark to observe due to the limits of condensing light performance and intensity of the light source, considering projection of whole astronomical objects in a planetarium. Therefore, to provide the experimental stimuli in our experiment, the hole diameter sizes of the star plates were designed considering all projection conditions, including intensity of light source, design filter, and optical conditions, and with the minimum requirement that the brightness of represented minimum magnitude stars could be perceived reliably. The relationship between the magnitude of each fixed star and its size projected onto the dome screen by the star plates is shown in Fig. 3(b). The projection size was categorized according to the star magnitude. For example, both stars of magnitudes 5.4 and 5.7 were reproduced with about 1 mm hole as shown in Fig. 3(b).

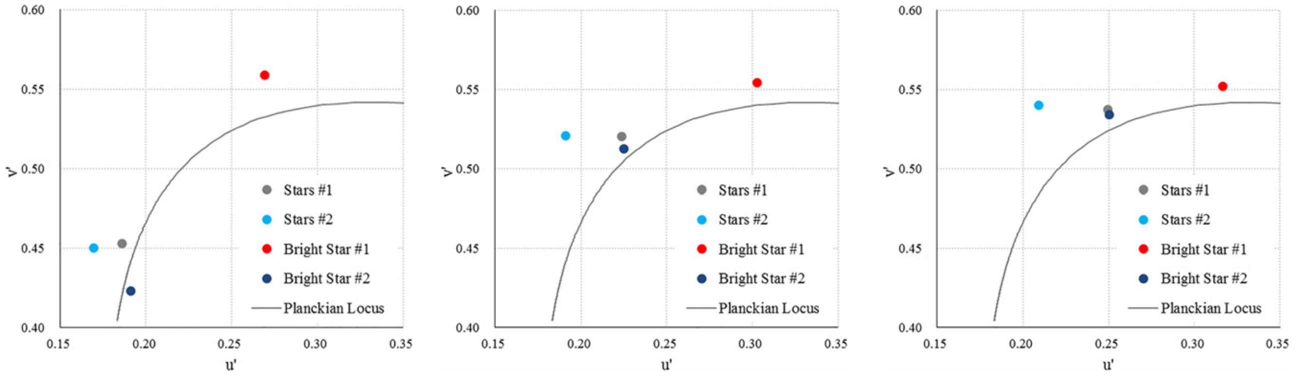
The projection size of stars was determined by magnitude class, which was categorized to present the equivalent of tone difference. The stars were projected in the range of $0.2'$–$4'$ of viewing angle from the viewing position of the observers. As explained above, bright stars such as Betelgeuse and Rigel were reproduced with near natural brightness using a different and higher-powered projector separated from the fixed star reproduction of the star plate. The color of these stars projected onto the dome screen was measured using a spectroradiometer (CS-2000, Konica Minolta). Due to limitation of the color-conversion filters, the chromaticity shown in Fig. 3(a) deviated from the original star color (based on the chromaticity of the Planckian locus); however, we confirmed that the projected chromaticity was adequately faithful to provide a perceived color near that of the actual star by experienced observers with enough experience in astronomical observation. In this confirmation process, the faithfulness was confirmed using the memory matching method on the dark dome of the planetarium. From the above procedures, we created and fixed a standard pattern.

Color-temperature shift pattern

When color is perceived using human mesopic vision, it has been reported that not only perception of saturation and intensity decline but also that there is a shift in the perception of hue.[12] We investigated faithful color reproduction to reveal the influence of star color as perceived in actual astronomical observation, by generating patterns to shift the color temperature of the whole star image. For the standard pattern with confirmed perceptually faithful reproduction, three versions of Pattern C were prepared with shifted color temperature. Pattern C-1, C-2, and C-3 had color-temperature shifts of −100 mired, +100 mired, and +200 mired, respectively, from the standard pattern using color-conversion filters. In addition to the patterns C-1 and C-2 (±100 mired), pattern C-3 (+200 mired shift with strong reddish color) was provided because the visual sensitivity of long wavelengths becomes weak due to the Purkinje effect in mesopic vision. Figure 4 shows the measured chromaticity values of the projected stars. It was confirmed that the color distance between Bright star #1 and the other three stars changed greatly. In these C-patterns, color was changed but luminance and size remained as in the standard pattern.

Luminance shift pattern

As a background to point light sources, the brightness of the night sky has a big influence on the perception of stars during actual observation. When the darkness of the



(a) (b) (c)

Figure 4. Chromaticity values of stars for each C-pattern in the CIE u′v′ chromaticity diagram. (a) C-1. (b) C-2. (c) C-3.

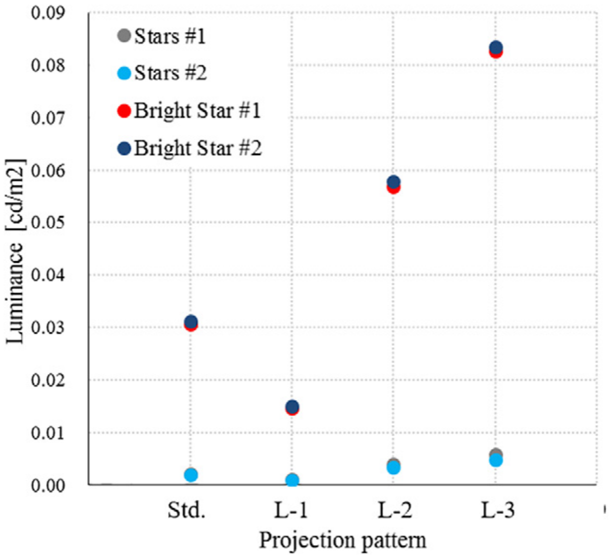


Figure 5. The luminance of the stars in each pattern.

night sky is altered by light pollution and zodiacal light, the perceptible star magnitude is limited. It seems that contrast between the brightness of the star and that of the night sky always influences perception of stars during astronomical observation. The changes in contrast between the brightness of the surrounding night sky and a star may increase or decrease its visibility and the feeling of the brilliance of the star. Therefore, we changed the contrast with the background of the night sky and visibility by preparing the L-patterns. For these, the projected luminance of the entire star field was relatively increased or decreased using ND filters. Figure 5 shows the luminance of stars in the standard pattern and the L-pattern shifts in luminance. As shown in Fig. 5, the luminance of the stars in pattern L-1, L-2, and L-3 were approximately 0.5, 1.9, and 2.7 times, respectively, the luminance of the standard pattern. These L-patterns changed luminance but retained the color and size of the standard pattern.

Size change pattern

In the actual starry sky, an individual star has an infinitesimal physical viewing angle. It is desirable to display a star image at its original size to provide a faithful reproduction of a starry sky. However, much brighter light sources would be necessary because the perceived brightness is reduced when the star size resulting from projection is smaller than normal. However, it is very difficult to reproduce images with the natural size and brightness of actual stars with the limitations of the projection system in a planetarium. Therefore, we controlled the size of the stars projected on dome screen by relatively expanding or contracting the hole diameter size of the star plate for the standard pattern. This allowed us to investigate the most appropriate projection size for reproducing faithfully a starry sky. Figure 6 shows the photographs of projected stars on the screen in the S-patterns, for which projection size had been changed, compared with the standard pattern. The luminance and color of the stars in the S-patterns were the same as those of the standard pattern, but their brightness as perceived by observers was made different from the standard pattern by changing the projection size. Compared with the standard pattern, patterns L-1 and L-2 had disc area rations 2/3 and 3/2, respectively.

Planckian locus pattern

Pattern C was only the pattern in which the color temperature was shifted from the standard pattern. Besides color shift, there is arbitrariness in the method of determining the reproduced color of stars, but we used a meaningful color determination method. We prepared Pattern B with the approximated chromaticity of the Planckian locus physically, because the actual star color had the chromaticity on its locus. Then we used the color index of a star catalog to design the star color of the experimental stimulus.[16,17] This Pattern B was implemented by superimposing the mixture color using a sharp-cut filter, a color-conversion filter, and an ND filter over the light source of the standard pattern. The measurement of the projected star color in Pattern B is shown in Figure 7. We can confirm that the chromaticity of stars in

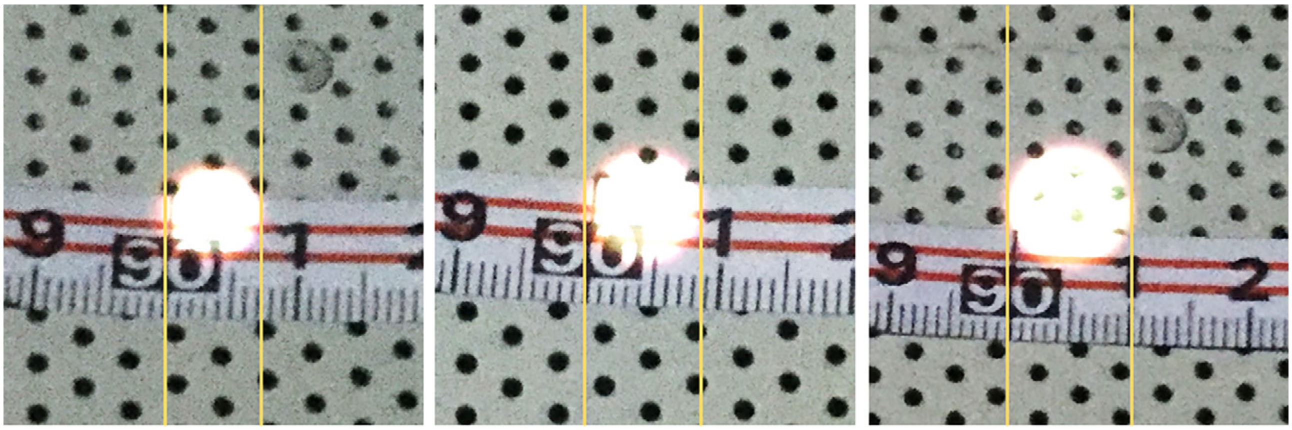


Figure 6. Size comparison of the star in Pattern S (from left to right, Pattern S-1, Pattern Std, Pattern S-2).

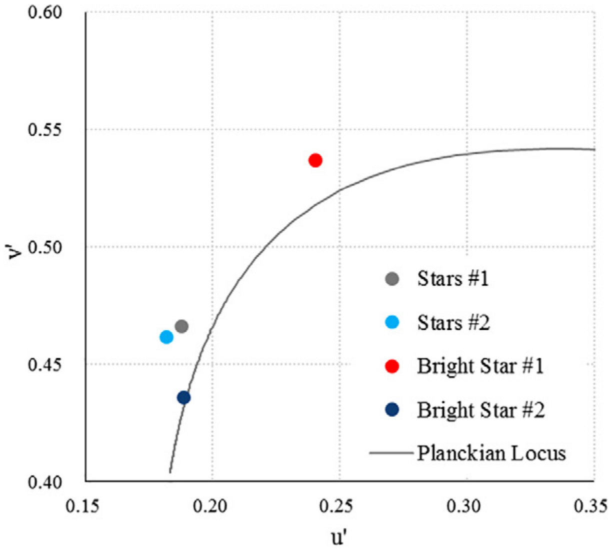


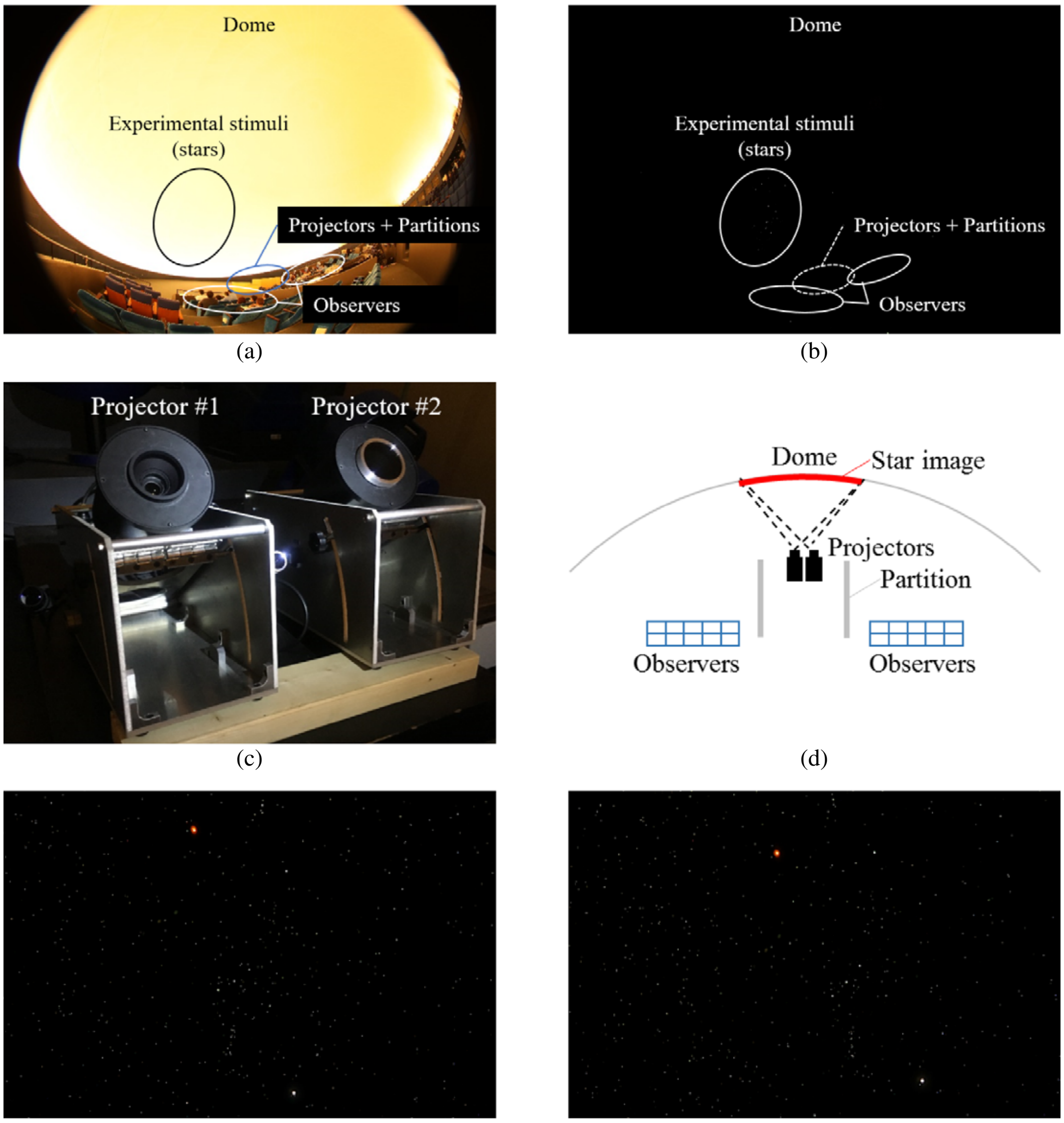
Figure 7. The star color for Pattern B in the CIE u′v′ chromaticity diagram.

Pattern B were closer to the color on the Planckian locus of actual stars than were the other patterns.

Experimental method

We conducted a psychophysical experiment to assess the perception of faithfulness of star image reproduction in a planetarium for ten projection patterns, as summarized in Table 1. The observers evaluated the results (compared with the actual starry sky) using opposite word pairs ("faithful"/"non-faithful") and five integer levels from −2 to +2, and wrote their evaluation values down on answer cards. The meanings of each evaluation level were −2 (not faithful), −1 (slightly not faithful), 0 (not which), +1 (slightly faithful), and +2 (faithful). The answer task was conducted in darkness with only the projected star image to maintain dark adaptation, but there was no other bias to discriminate against particular answers. In the evaluation, there was no designated fixation point and the observers were able to observe the star image freely. Therefore, they could judge the color and brightness of the whole projection stimuli via the foveal vision by the cones and the peripheral vision by the rods well. Snapshot images of the experimental environment are shown in Figure 8. Fig. 8(a) shows an image of the illuminated dome captured using a fish-eye lens and Fig. 8(b) shows an image of the dark dome captured under the same photographic conditions. Each star pattern was projected to the position of the oval mark in each figure. The diameter of dome screen was 23 m and the zenith of the dome screen was slanted 15° to the front. There was no other illumination in the space where the experiments were conducted, aside from that of the projected starry sky image. It was not possible to verify the low light level using the CS-2000 because it was too dark to measure (<0.003 cd m^{-2}). The room appeared completely dark. The averaged viewing distance between the observers and the center of projected star field image was 10.7 m, and the distance slightly varied according to the seat position (max: 11.9 m, min: 9.6 m). The viewing angle of the projected star image at the averaged viewing distance was 37.3°. The distance between the dome screen and the projectors (Fig. 8c) was 7.75 m. The projectors were surrounded with partitions that prevented the leakage from the light source reaching the observers, as shown in Fig. 8(d), in which it can also be seen that the degree of leaning of the perceived star image was slightly different because the observers sat to the right or left of center. We show the experimental stimuli (star field image) that observers could see from the right side and the left side in Fig. 8(e) and (f). All observers had cushion reclining seats and observed the projected star image in front of their eyes in a relaxed posture. The room temperature in the dome was 26°C and the conditions remained comfortable and uniform for observation throughout the experiment.

Twenty observers, including eighteen men and two women, in their twenties to sixties participated in this experiment, as shown in Figure 9. All the observers had experience making astronomical observations of an actual starry sky. Therefore, the observers could evaluate the faithfulness of the projected star image by memory matching comparison for the starry sky around Orion which was memorized by each observer. This psychophysical experiment was conducted



(a) (b)

(c) (d)

(e) (f)

Figure 8. Snapshot of the experimental environment. (a) Captured in bright dome. (b) Captured in dark dome. (c) Projectors for fixed stars. (d) Observers' position in the planetarium. (e) Experimental stimulus for left side observers. (f) Experimental stimulus for right side observers.

using the dome of the planetarium. After the observers taking a seat in the dome, the illumination in the dome was turned off. We confirmed the brightness in the dark dome was less than magnitude 23 using the Sky Quality Meter. Most stars with magnitude more than 7.4 reproduced by projection could be perceived because the magnitude limit for observation in the darkness of the 23.0 magnitude was about magnitude 7.0.[14]

At the beginning of this psychophysical experiment, the observer received instructions for the evaluation experiment and did exercises using a practice pattern for 20 minutes. It was assumed that the observers had completed dark adaptation by this time. There was no illumination except from the projected star field image, and the experiment was preceded by oral instructions using a microphone in the dark dome. In this experiment, 11 randomly projected patterns (including standard patterns twice) were used to confirm reproducibility, and the observers evaluated the faithfulness of each star field image (i.e., each pattern) within 15 seconds after observing the star image for 30 seconds. We set 30

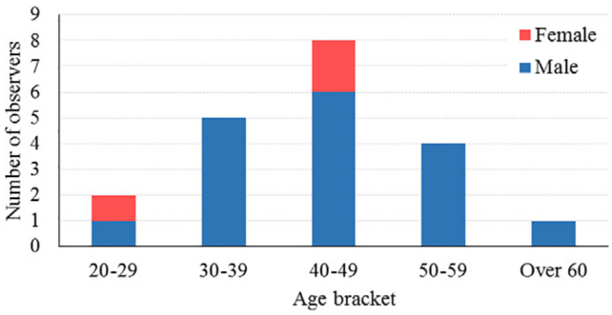
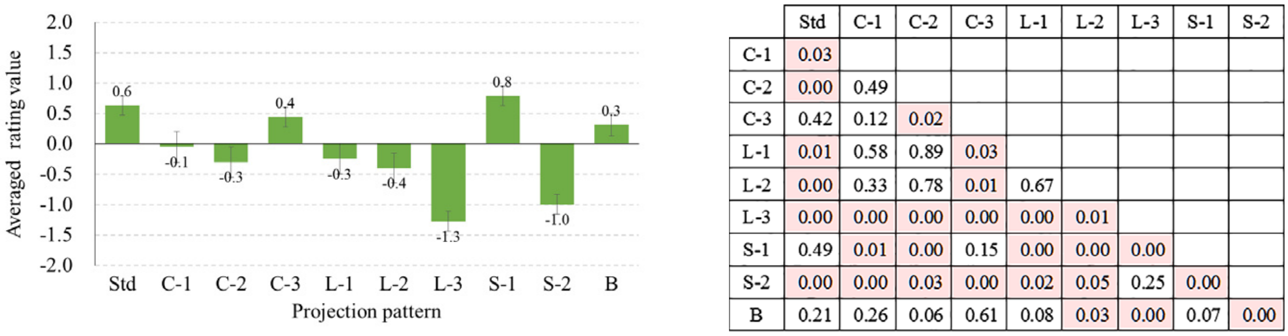


Figure 9. Number of observers of each age.



	Std	C-1	C-2	C-3	L-1	L-2	L-3	S-1	S-2
C-1	0.03								
C-2	0.00	0.49							
C-3	0.42	0.12	0.02						
L-1	0.01	0.58	0.89	0.03					
L-2	0.00	0.33	0.78	0.01	0.67				
L-3	0.00	0.00	0.00	0.00	0.00	0.01			
S-1	0.49	0.01	0.00	0.15	0.00	0.00	0.00		
S-2	0.00	0.00	0.03	0.00	0.02	0.05	0.25	0.00	
B	0.21	0.26	0.06	0.61	0.08	0.03	0.00	0.07	0.00

(a) (b)

Figure 10. Evaluation result of all observers for each projection pattern. (a) Averaged rating value. (b) *p*-value of between-pattern *T*-tests.

seconds for observation because Blackwell reported that detection limit of the brightness in the dark environment becomes constant after more than 15 seconds in the experiments.[18] Between each pattern projection, observers had a short break of several minutes while the projection pattern was reset. The observers were asked to observe the star field image of the standard pattern projected between the short breaks. This caused resetting of the influence of the anteroposterior evaluation pattern on the observers' perception. As shown in Fig. 8(c), two projectors were used to change the projections smoothly: one was used to project the evaluation patterns and the other one was always used to project the standard pattern. A break time of about 10 minutes was also taken to refresh the observers and the experiment was started again after dark adaptation. The illumination of the dome was turned on after all evaluation tasks related to this experiment were finished. Then observers answered a questionnaire in the lit place and left the room.

Experimental results

The significance of the answered evaluations for each pattern was verified using a t-test after excluding the outlier data using the Smirnoff–Grubbs test. The distribution of evaluations for each projection pattern was also checked by F-test to see whether the variances were equal. There was no significant difference in the answer results for the standard pattern, which was evaluated two times for confirmation of reproducibility. Therefore, we used the second result of the standard pattern for the analysis. The intra-observer variance calculated from evaluation of two passes with the standard patterns was 0.34, and the inter-observer variance calculated from evaluations of all projected patterns was 0.80.

Total averaged rating value

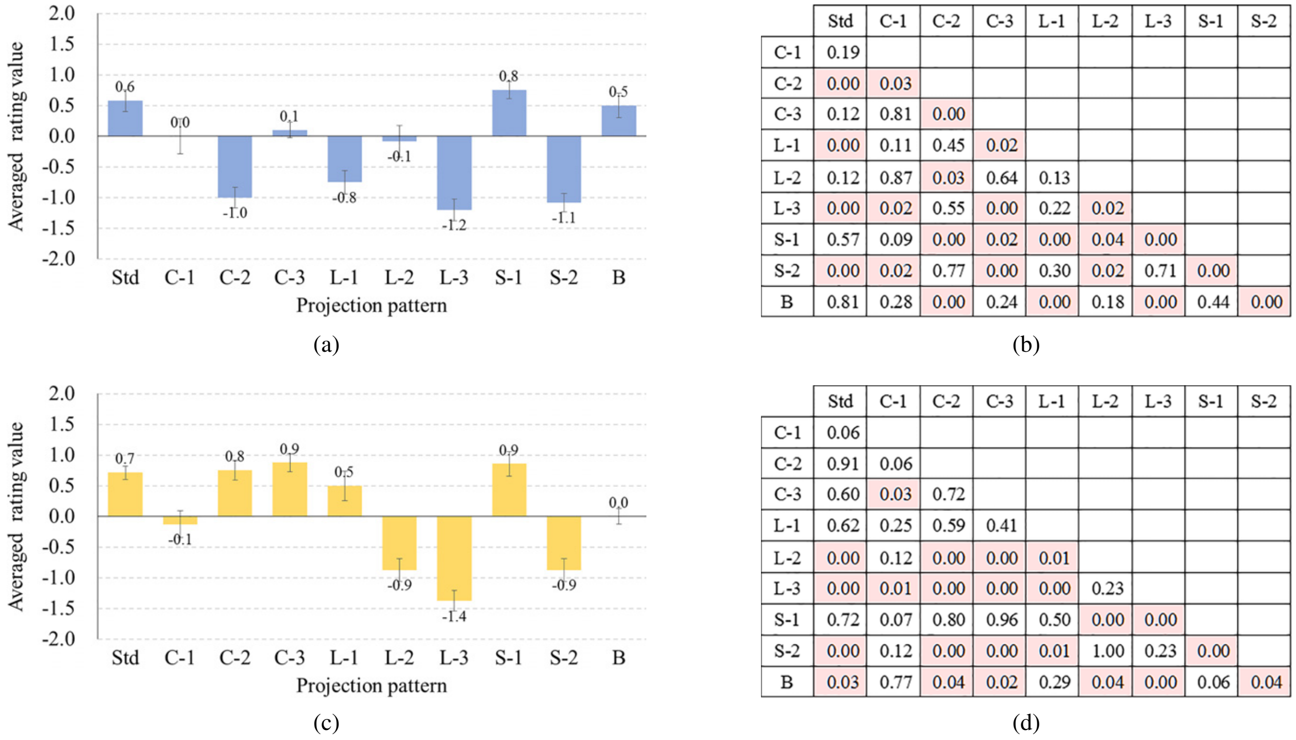
We show the averaged rating value (with the standard error among all twenty observers) for each pattern in Figure 10(a). Fig. 10(b) also shows a significant difference in the answer rating for each projection pattern. Each rating shows the p-value of between-pattern T-tests. Color shows when a p-value indicated a significant difference ($p < 0.05$).

The projection pattern with the highest rating of faithfulness was pattern S-1, as shown in Fig. 10(a). The significant difference between pattern S-1 and the other patterns show that there was no significant difference between the four patterns with high rating values (shown in Fig. 10b). In other words, the evaluations of faithfulness were significant for the projection patterns S-1, standard, C-3 and B, in the order of higher faithfulness. On the other hand, the projection pattern with the lowest faithfulness rating was L-3. The significant difference with pattern L-3 shows that the projection patterns L-3 and S-2 were evaluated as significantly not faithful, as shown in Fig. 10(b). From these results, projection pattern L-3 with high intensity, and Pattern S-2 with high perceptual brightness (caused by enlarging the projection size) had predominantly negative evaluations for faithfulness. These results suggest that luminance and projection size greatly influenced the faithfulness evaluation. In particular, all four of the projection patterns evaluated as having high faithfulness had the same luminance as the standard pattern. It appears from these results that the observers were sensitive to change in luminance during the faithfulness evaluations. In contrast, the tolerance to lack of faithfulness was high regarding the color-temperature shift in the prepared patterns used in this experiment. A detailed consideration of each pattern is shown in a later section.

Classification of observers

In the analysis of the rating scores provided by observers, a specific tendency was found for each image pattern. We classified the observers based on the rating scores of all twenty observers for each pattern, using hierarchical clustering. As the result, observers were sorted into two groups (Group 1 and Group 2). The number of observers in Group 1 was thirteen, and in Group 2 was seven. According to their answers on the questionnaire, the observers in Group 1 were mainly in their 30s–40s, while those in Group 2 were mainly in their 50s–60s. There was no difference between observers reflecting their observation experience.

The average rating values for each group for each projection pattern are shown in Figure 11(a) and (c). Figs. 11(b) and (d) show the significant difference between each projection pattern. As shown in Fig. 11(a) and (c),



(a)

	Std	C-1	C-2	C-3	L-1	L-2	L-3	S-1	S-2
C-1	0.19								
C-2	0.00	0.03							
C-3	0.12	0.81	0.00						
L-1	0.00	0.11	0.45	0.02					
L-2	0.12	0.87	0.03	0.64	0.13				
L-3	0.00	0.02	0.55	0.00	0.22	0.02			
S-1	0.57	0.09	0.00	0.02	0.00	0.04	0.00		
S-2	0.00	0.02	0.77	0.00	0.30	0.02	0.71	0.00	
B	0.81	0.28	0.00	0.24	0.00	0.18	0.00	0.44	0.00

(b)

(c)

	Std	C-1	C-2	C-3	L-1	L-2	L-3	S-1	S-2
C-1	0.06								
C-2	0.91	0.06							
C-3	0.60	0.03	0.72						
L-1	0.62	0.25	0.59	0.41					
L-2	0.00	0.12	0.00	0.00	0.01				
L-3	0.00	0.01	0.00	0.00	0.00	0.23			
S-1	0.72	0.07	0.80	0.96	0.50	0.00	0.00		
S-2	0.00	0.12	0.00	0.00	0.01	1.00	0.23	0.00	
B	0.03	0.77	0.04	0.02	0.29	0.04	0.00	0.06	0.04

(d)

Figure 11. Evaluation result of each Group for each projection pattern. (a) Average rating value in Group 1. (b) *p*-value of between-pattern tests of ratings by Group 1. (c) Average rating value in Group 2. (d) *p*-value of between-pattern tests of ratings by Group 2.

the rating values of both groups were in accord in many patterns, but we confirmed that the evaluations of pattern C-2, C-3, and L-1 showed significant differences between groups ($p < 0.05$). The rating values between patterns C-2 and L-1, had a particularly large difference (i.e., opposite rating directions). Group 2 gave a high positive rating for each projection pattern, but Group 1 gave remarkably negative ratings. These results revealed that the evaluation of the faithfulness of the projection patterns C-2 and C-3 (low color-temperature shift) and L-1 (low luminance) involved large individual differences.

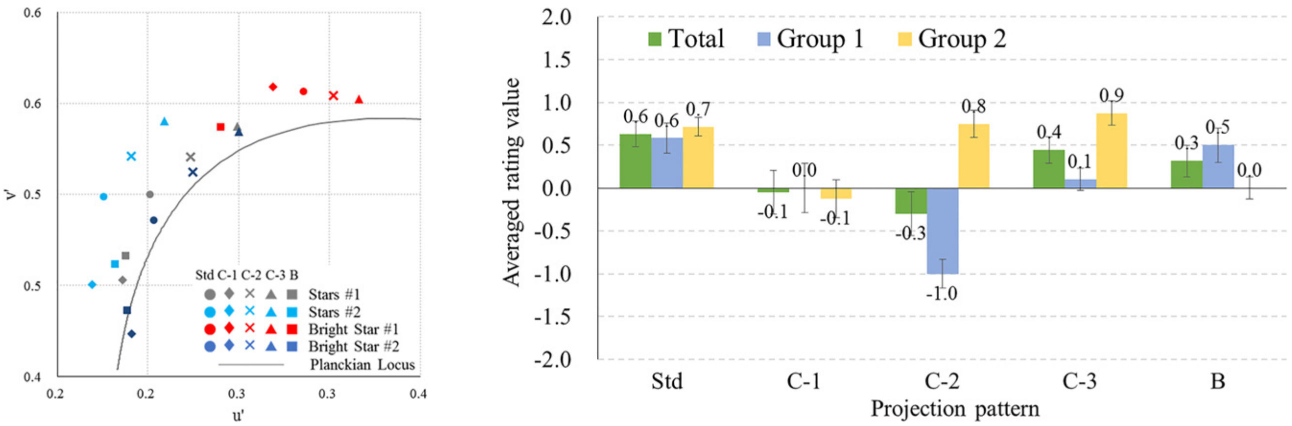
Important factors for faithful representation

In our psychophysical experiment, the faithfulness of the representation of star images with one changed parameter and two fixed parameters (among color, luminance, or size of projected stars) projected in a planetarium was evaluated. In this subsection, we consider the factors that influenced faithful reproduction of the starry sky for each parameter.

Color effect

We considered the color effect in this evaluation of the faithfulness of a projected star image by comparing the results for the standard pattern and the color controlled patterns (C-1, C-2, and C-3), with color temperature shifted from the standard pattern, and Pattern B which approached the color on the Planckian locus. Every pattern had the same luminance and size as the standard pattern.

In Figure 12, we show the difference in color among patterns and the ratings for these patterns. The error bars show the standard error in Fig. 12(b). Considering the overall rating of all the observers, it was positively judged to be faithful in the descending order: standard pattern, Pattern C-3 (shifted color temperature of +200 mired), and Pattern B (color approaching that on the Planckian locus). However, there was no significant difference among these patterns. Furthermore, all the patterns (C) with shifted color temperature did not get considerably low ratings (evaluation). These results show that the tolerance to lack of faithfulness is high for color changes. This new finding might be explicitly supported by a visual characteristic[19] that the sense of color difference becomes weaker when the viewing angle of the target object is small. Nevertheless, it was interesting that patterns C-3 and B got high ratings. This result seems to be attributable to the perception that Betelgeuse (which had the most saturated red color of all the projected stars) affected the judgement of faithful reproduction of the colors of other stars. The spectral distribution of real Betelgeuse has the shape of a soaring arc in the visible light region, from short wavelengths (400 nm) to long wavelengths.[17] The decline of reddish color perception in mesopic vision seems to be milder than was the perception off the LED light (restricted spectral distribution) used in our experiment with the projection system in the planetarium. However, luminance in the experimental environment was less than the measurement limit of 0.003 cd/m^2 of our instrument. This is considerably darker than the brightness of an actual starry sky. Therefore, the reddish reproduction of Pattern C-3 (+200 mired shift) must have been highly evaluated due to promotion of the decline in



(a) (b)

Figure 12. Standard and color change patterns. (a) Chromaticity of projected star color. (b) Rating value.

sensitivity to long wavelengths in the darkness. Furthermore, the enhancement of saturation due to the memory matching effect seems to be affecting to assessment. The main difference in the chromaticity of pattern C-1 and B, was the color of Betelgeuse, as shown in Fig. 12(a). Nevertheless, a significant difference between the two patterns was observed. This result suggests that the observers based their evaluation of pattern B on the entire color balance of the projected star field image.

Next, we focus on the rating scores in each group. The rating of Pattern C-2 by Group 1 was noticeably low, whereas the ratings of patterns C-2 and C-3 by Group 2 were high. A number of factors were considered as causes: individual differences in observer's perception, including the psychological factor of the difference in the strength to perceive the chroma; and physiological factors such as difference in visual sensitivity and eye pupil size. As for the age bracket of each group, those in Group 1 were in their 30s–40s, and those in Group 2 were in their 50s–60s (years old). The decline of color perception ability with age was considered the cause of this result because Group 2 positively evaluated (with high ratings) patterns C-2 and C-3, which exhibited increased red from low color temperature. Generally, color perception becomes weaker with decline in the illumination level, and the degree of the decline in perception has great individual difference.[20] The perceived retinal illuminance seems to reflect differences in the pupil size of individual eyes. Our understanding is that Group 2 (higher age bracket) positively evaluated Pattern C-2 (strong reddish color) because the maximum pupil size becomes smaller with aging.

Luminance effect

We considered the effect of luminance to influence the faithfulness evaluation. Figure 13 shows the rating result for projection pattern L (shifted luminance) in relation to the standard pattern. Every pattern had the same color and projection size as the standard pattern. In comparison with the standard pattern, the faithfulness evaluations for all L-patterns (L-1, L-2, and L-3) were lower. The factors that we expected to affect projection of the L-patterns, including design of projection conditions (such as the increase and decrease of the contrast between the stars and the background night sky) and change in visibility, were bad influences on judgement of the reproduction faithfulness. This result indicates that these patterns were evaluated negatively if the luminance of star image was darker (pattern L-1) or brighter (patterns L-2 and L-3) than the standard pattern, and the standard pattern had the best luminance for faithful reproduction of star field images in a planetarium. This means that the observers were very sensitive to change in luminance in perceiving the faithfulness of a starry sky, within the range of luminance in this experiment.

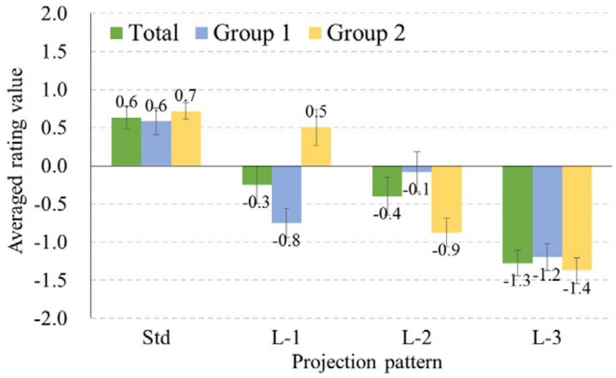


Figure 13. Rating value for the standard and luminance change patterns.

Size effect

The star size reproduced in the planetarium is about 0.2′–4′ of viewing angle, whereas the actual star size is perceived as approximately 1′ in the case the visual acuity of normal observers. The size reproduction is restricted by the projection environment and the system specifications in the planetarium. Even so, we investigated the influence on the evaluation of faithful reproduction by preparing the projection patterns S-1 and S-2 (changed projection size). Each pattern had the same color and luminance of stars as the standard pattern. Figure 14 shows the ratings for the standard pattern, and for patterns S-1 and S-2 with resized hole

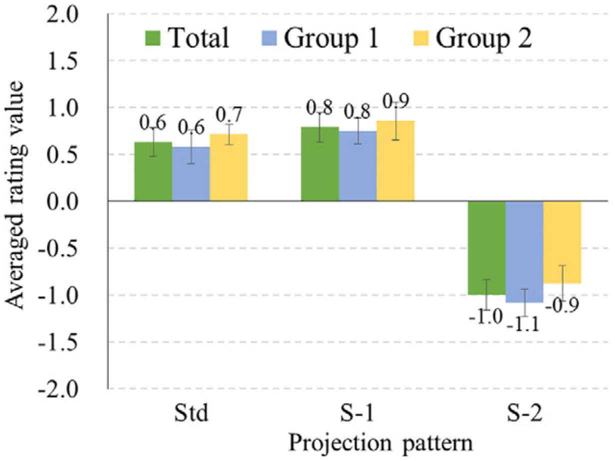


Figure 14. The rating value for the standard and projection size change patterns.

diameter. As expected, the projection pattern S-1 (smaller projection size of 2/3 the area ratio) was positively evaluated (better than the standard pattern). On the other hand, the projection pattern S-2 which had enlarged projection size (3/2 the area ratio) was given a considerably negative evaluation. It is desirable that the projection size be smaller than the projection size of the standard pattern, at least for faithful reproduction of star images because there was no significant difference between the ratings of the standard pattern and pattern S-1.

On the other hand, we could expect that it is hard to perceive darker stars after contraction of the projection size. There were observers who replied in a questionnaire that the standard pattern offered a different impression because more dim stars could be perceived than in an actual starry sky. The smaller projection size has the potential to improve the faithfulness of reproduction, but the visibility decreases at the same time. The dim stars were visible because the brightness of the night sky (e.g., zodiacal light and light pollution) was reduced in our experiment. However, in actual planetarium projections of illuminated environments involving moonlight, the reproduction of dim stars may not be perceived. For faithful reproduction of a starry sky in a planetarium, it is essential to establish projection technology able to duplicate the actual brightness of a star.

CONCLUSIONS

A natural starry sky is a set of point light sources, and a reproduction of stars is a special object beyond the categories of conventional vision studies for natural images and patches. In this study, to investigate the factors required to reproduce star field images in a planetarium faithful to an actual starry sky, we analyzed the faithful reproduction of stars by conducting psychophysical experiments using the projected stars as experimental stimuli, and changing three kinds of parameters (color, luminance, and size). A standard projection pattern for reference was designed by a group of experienced observers with abundant astronomical observation experience. The standard was faithful to the actual starry sky perceptually, but not physically. Twenty observers with astronomical observation experience conducted the evaluation experiments for ten kinds of star image patterns projected on a dome screen of 23 m diameter. As result, we revealed that darker or brighter reproduction than the standard reference pattern spoiled the faithfulness because the luminance of the stars was sensitive to faithful reproduction. In addition, in the faithfulness evaluation for the projection pattern with smaller star size, the rating of faithfulness improved. In contrast, the rating was remarkably negative when the pattern had stars of larger size than in the standard pattern. For shifts in the color temperature of stars, the observers could distinguish differences in color but did not give negative ratings for changes in the color pattern. These results indicated that the tolerance was high for lack of faithfulness in the color. With the color change pattern, positive ratings were given for increased reddish color, and the entire balance of the color approached that on the Planckian locus.

Because the projection range was limited to 1/32 the size of the whole sky in this experiment, there were patterns that could be realized by projection of the star field image. As an example, it is difficult to project a star pattern with size changed to 2/3 the area ratio of the whole sky due to technical issues such as the techniques available for processing star plates and the need for excellent optical performance to reproduce bright and dim stars clearly. By resolving these issues, we plan to investigate how to provide faithful reproduction of the whole starry sky in the future.

REFERENCES

[1] International Astronomical Union, http://www.iau.org/.

[2] International Dark-Sky Association, http://www.darksky.org/light-pollution/.

[3] International Planetarium Society, http://www.ips-planetarium.org/.

[4] G. A. Good, "Observing variable stars," *Springer Science and Business Media* (Springer-Verlag, London, UK, 2012).

[5] H. W. Jensen, F. Durand, M. M. Stark, S. Premože, J. Dorsey, and P. Shirley, "A physically-based night sky model," *Proc. 28th Aannual Conf. on Computer Graphics and Interactive Techniques* (ACM, New York, NY, USA, 2001), pp. 399–408.

[6] H. W. Jensen, S. Premože, P. Shirley, W. B. Thompson, J. A. Ferwerda, and M. M. Stark, "Night rendering," *Technical Report UUCS-00-016* (2000).

[7] J. A. S. Kinney, "Comparison of scotopic, mesopic, and photopic spectral sensitivity curves," J. Opt. Soc. Am. **48**, 185–190 (1958).

[8] A. G. Kirk and J. F. O'Brien, "Perceptually based tone mapping for low-light conditions," ACM SIGGRAPH **42**, 1–10 (2011).

[9] M. Mikamo, B. Raytchev, T. Tamaki, and K. Kaneda, "A tone reproduction operator for all luminance ranges considering human color perception," *EUROGRAPHICS, 2014* (The Eurographics Association and John Wiley & Sons, Ltd., Hobokeny, NJ, USA) .

[10] J. C. Shin, N. Matsuki, H. Yaguchi, and S. Shioiri, "A color appearance model applicable in mesopic vision," Opt. Rev. **11**, 272–278 (2004).

[11] J. Pokorny, M. Lutze, D. Cao, and A. J. Zele, "The color of night: surface color perception under dim illuminations," Vis. Neurosc. **23**, 525–530 (2006).

[12] J. C. Shin, H. Yaguchi, and S. Shioiri, "Change of color appearance in photopic, mesopic and scotopic vision," Opt. Rev. **11**, 265–271 (2004).

[13] C. Leinert, S. Bowyer, L. K. Haikala, M. S. Hanner, M. G. Hauser, A. C. Levasseur-Regourd, I. Mann, K. Mattila, W. T. Reach, W. Schlosser, H. J. Staude, G. N. Toller, J. L. Weiland, J. L. Weinberg, and A. N. Witt, "The 1997 reference of diffuse night sky brightness," Astron. Astrophys. Suppl. Ser. **127**, 1–99 (1998).

[14] B. E. Schaefer, "Telescopic limiting magnitudes," Astron. Soc. Pac. **102**, 212–229 (1990).

[15] D. Jones, "Norman pogson and the definition of stellar magnitude," Astron. Soc. Pac. Leaflets **10**, 145–152 (1967).

[16] M. A. C. Perryman, L. Lindergren, L. Kovalevsky, E. Hoeg, U. Bastian, P. L. Bernacca, M. Crézé, F. Donati, M. Grenon, M. Grewing, F. can Leeuwen, H. van der Marel, F. Mignard, C. A. Murray, R. S. Le Poole, H. Schrijver, C. Turon, F. Arenou, M. Froeschlé, and C. S. Petersen, "The HIPPARCOS Catalogue," Astron. Astrophys. **323**, L49–52 (1997).

[17] G. H. Jacoby, D. A. Hunter, and C. A. Christian, "A Library of Stellar Spectra," Astrophys. J. Suppl. Ser. **56**, 257–281 (1984).

[18] H. R. Blackwell, "Contrast Thresholds of the Human Eye," J. Opt. Soc. Am. **36**, 624–643 (1946).

[19] R. C. Carter and L. D. Silverstein, "Size matters: Improved color-difference estimation for small visual targets," J. Soc. Inf. Disp. **18**, 17–28 (2010).

[20] C. Momma, S. Honma, H. Yaguchi, H. Haneishi, and Y. Miyake, "Color appearance and color reproduction for mesopic vision," J. Opt. Soc. Am. **4**, 161–163 (1992).