

Understanding individual differences in color appearance of “#TheDress” based on the optimal color hypothesis

Research Institute for Multimodal Sensory Science,
Faculty of Human Sciences, Kanagawa University,
Yokohama, Japan

Keiji Uchikawa

Department of Information Processing,
Tokyo Institute of Technology, Yokohama, Japan



Takuma Morimoto

Department of Information Processing,
Tokyo Institute of Technology, Yokohama, Japan



Tomohisa Matsumoto

Research Institute for Multimodal Sensory Science,
Faculty of Human Sciences, Kanagawa University,
Yokohama, Japan



We performed a theoretical analysis based on our optimal color hypothesis to explain why “#TheDress” image had a different color appearance for different observers (observer-dependent perception). We then carried out an experiment to test the hypothesis derived from the aforementioned theoretical analysis. In the optimal color hypothesis, the visual system picks the optimal color distribution that provides the best fit to the luminance distribution at a scene. The peak of the best-fit optimal color distribution corresponds with the illuminant’s color temperature. In the theoretical analysis, we found that as the luminance level was increased the best-fit optimal color temperature changes abruptly from high to low at a specific luminance-level. Under the dark-blue (low luminance and high color temperature) illuminant the dress should appear white/gold whereas under the bright-white (high luminance and low color temperature) illuminant the dress should appear blue/black. The observer-dependent appearances of the dress may be explained by this luminance-dependent illuminant prediction. In the experiment, we used the original dress, a chromatically inverted dress, and an achromatic dress as stimuli. The observer adjusted chromaticity and luminance of a test field drawn onto the dress image so that it appeared as a full-white surface. We found that the white/gold group estimated the illuminant to be darker and bluish and the blue/black group estimated it to be brighter and yellowish. The observer’s estimated illuminant was consistent with the predicted illuminant by the optimal color theory. It was newly discovered that even when the dress was achromatic, these two

groups estimated the illuminant to be darker or brighter in the same way as for the original dress.

Introduction

In February 2015, the image of “#TheDress” posted on the internet spread quickly across the world. The image of “#TheDress” attracted unusually strong attention from people because judgments about its color appearance were mostly divided into two groups depending on the individual observers: white/gold or blue/black. This discrepancy was very clear among people unlike an ordinal color illusion pattern, which caused the same illusion in most people (Brainard & Hurlbert, 2015).

Color scientists recognized this problem as a novel phenomenon since we have not previously encountered a stimulus with such clear individually-divided color appearances. Quite a few studies have been reported on the dress phenomenon since 2015 to try to explain individually-divided color appearances based on current color vision theories (Brainard & Hurlbert, 2015; Gegenfurtner, Bloj, & Toscani, 2015; Lafer-Sousa, Hermann, & Conway, 2015; Schlaffke et al., 2016; Winkler, Spillmann, Werner, & Webster, 2015) and to reveal any imaging and viewing factors relevant to the dress appearance (Hesslinger & Carbon, 2016; Melgosa, Gómez-Robledo, Suero, & Fairchild, 2015; Vemuri, Bisla, Mulpuru, & Varadharajan, 2016). However,

Citation: Uchikawa, K., Morimoto, T., & Matsumoto, T. (2017). Understanding individual differences in color appearance of “#TheDress” based on the optimal color hypothesis. *Journal of Vision*, 17(8):10, 1–14, doi:10.1167/17.8.10.

doi: 10.1167/17.8.10

Received September 30, 2016; published July 28, 2017

ISSN 1534-7362 Copyright 2017 The Authors



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Downloaded From: <http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/936361/> on 09/05/2017

currently no one has been able to fully explain this phenomenon.

Gegenfurtner et al. (2015) measured the colors of the cloth and the lace parts in the dress by color matching it with an adjustable circle adjustment stimulus presented next to the dress image on an LCD monitor. They found that the chromaticities matched to the dress image were overlapping, but the luminances were well separated between the two groups. Lafer-Sousa et al. (2015) showed three peaks corresponding with white/gold, blue/black and blue/brown by matching the colors of the dress. This suggested that different brains resolve the dress image into distinct stable percept. Winkler et al. (2015) pointed out that the visual system has a blue–yellow asymmetry to perceive surface colors as white or gray. They showed that observers were more likely to use “white” to name colors in a bluish direction than those in a yellowish direction (Churma, 1994; Pearce, Crichton, Mackiewicz, Finlayson, & Hurlbert, 2014). Brainard and Hurlbert (2015) stated that the authors of these three studies invoked explanations related to color constancy (Foster, 2011; Hurlbert, 2007), and described a plausible explanation that people who perceive a blueish illumination in the dress image see the dress as white/gold, and those who perceive a yellowish illumination see the dress as blue/black.

Most studies emphasized that how observers perceived the illumination in the dress image would be a crucial factor. This illuminant estimation by observers should be considered in resolving the dress color ambiguity. We recently developed the optimal color hypothesis to explain color constancy phenomena (Fukuda & Uchikawa, 2014; Morimoto, Fukuda, & Uchikawa, 2016; Uchikawa, Fukuda, Kitazawa, & MacLeod, 2012). In the optimal color hypothesis, the visual system estimates an illuminant in a scene (with a given luminance-chromaticity distribution) by using an optimal color shell of the illuminant’s color temperature. In the present study, we tried to explain the dress’s color appearances using our optimal color hypothesis. In the first section, we applied the optimal color hypothesis to the dress image to elicit two categories of color appearance (white/gold and blue/black) and demonstrate the individual differences between the two categories. In the second section, we carried out an experiment of illuminant estimation for the dress image to test whether the illuminant could be estimated to be dark and bluish by the white/gold group, and to be bright and yellowish by the blue/black group. It was demonstrated, theoretically and experimentally, that the optimal color hypothesis could provide an explanation for at least two possible appearances of the dress image.

Theoretical analysis

Optimal colors

The optimal color is defined as a surface having the maximum luminance attainable at a given chromaticity under a given illuminant. It has two abrupt spectral transitions between 0 and 100% reflectance. Figure 1a shows the luminance distribution of optimal colors under 3000K, 6500K, and 20000K black-body illuminants on a MacLeod–Boynton (M–B) chromaticity diagram. We used the Stockman and Sharpe (2000) spectral sensitivity of L, M, and S cones to calculate the M–B chromaticity coordinates and luminance. The optimal color stimuli were generated by incrementing each of the two spectral transition wavelengths in 4 nm steps and plotting a point for each optimal color in Figures 1a, b, and c. The optimal color forms the outer shell of the luminance distribution, covering all real surface colors in a luminance and chromaticity space. The shape of the outer shell changes depending on the illuminant.

Figure 1b shows the profiles of the optimal colors on the redness axis, $L / (L + M)$. The white surface, having 100% reflectance across all wavelengths, has the maximum luminance. However, a nonwhite surface, for example a red surface having no reflectance in shorter wavelengths, has lower luminance. The luminance distribution of optimal colors becomes triangular in shape. When the illuminant changes from 20000K to 3000K, the luminance distribution inclines toward the redness direction with the shift in the peak of the luminance distribution, which corresponds with the white surface with 100% reflectance. Surfaces of higher luminance shift more than those of lower luminance. The peak of the optimal color shell always corresponds with the chromaticity of the illuminant.

All surfaces existing in the real world are inevitably included in the optimal color shell. In Figure 1c, the dots represent 574 natural objects that Brown (2003) collected and the surfaces from 16 hyper spectral images measured by Foster et al. (2006). The envelope of luminance distribution of these natural colors closely resembles the optimal color shell. In this case the illuminant is 6500K. The natural color objects will change their positions within the optimal color shell under various illuminants

Optimal color hypothesis

In our optimal color hypothesis, it is assumed that the visual system takes into account the luminance distribution of a scene, then selects an optimal color shell that best fits to the luminance distribution. The

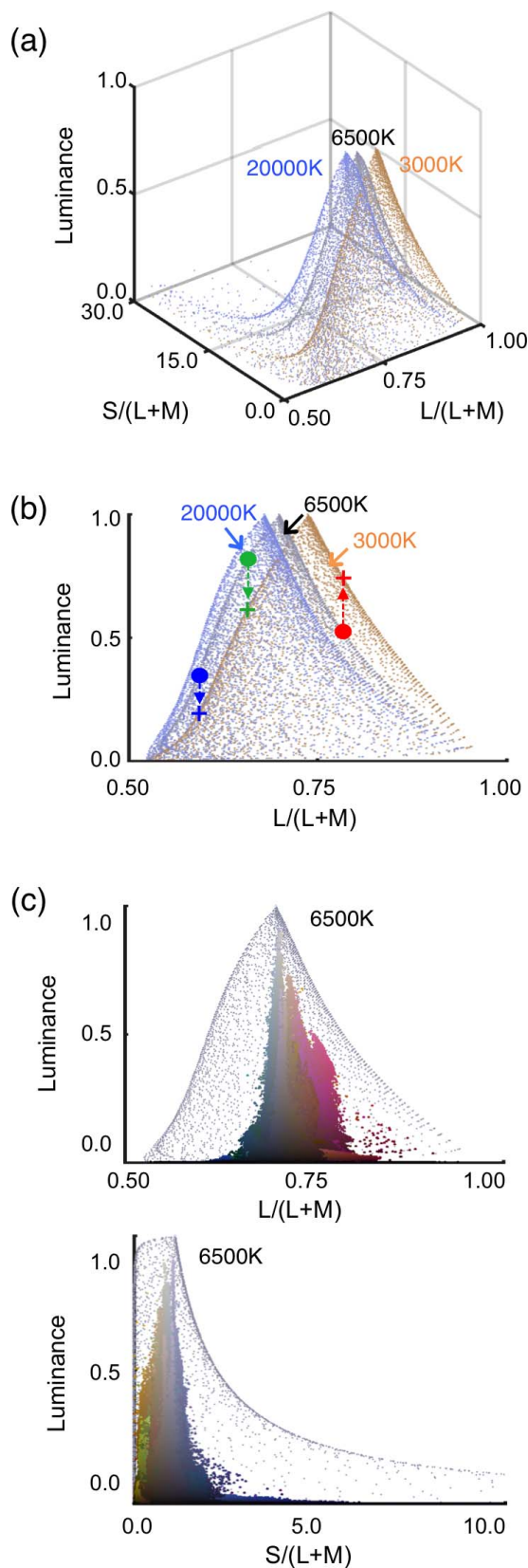


Figure 1. (a) The luminance distribution of optimal colors under black-body illuminants 3000K, 6500K, and 20000K on a MacLeod–Boynton chromaticity diagram. (b) The profiles of the optimal colors on the redness axis, $L/(L+M)$. Red, green

and blue circles represent surface examples in a scene (see text for details). (c) 574 natural objects (Brown, 2003) and surfaces in 16 hyperspectral images (Foster et al., 2006) plotted in the optimal color shell of 6500K. Colors of dots for natural objects and surfaces roughly correspond with color-appearance of those regions in the M–B chromaticity diagram.

peak of the selected optimal color shell corresponds with the chromaticity and the luminance of the estimated illuminant. For examples in Figure 1b, red, green, and blue circles represent three surfaces in a scene. All three circles are located on the limit of the optimal color shell of 6500K, but the red circle exceeds the limit of 20000K, and the green and blue circles exceed the limit of 3000K. In this luminance profile of surfaces, the optimal color shell of 6500K turns out to be the best-fit shell. Now, when the luminances of the green and blue surfaces decrease to the green and blue cross positions, and the luminance of the red surface increases to the red cross position with no chromaticity change, the 6500K optimal color shell no longer remains the best fit shell, but instead the 3000K optimal color shell takes its place. It has been shown that this hypothesis works quite well in various stimulus conditions (Fukuda & Uchikawa, 2014; Morimoto, Fukuda, & Uchikawa, 2016; Uchikawa, Fukuda, Kitazawa, & MacLeod, 2012).

We applied the optimal color hypothesis to the dress phenomenon. To begin, to clarify color distribution of the dress image, we calculated chromaticity and luminance of all pixels in the dress image, which were sampled in the area surrounded by back solid lines, shown in Figure 2a. The small black dots in Figure 2b represent the chromaticities of the pixels in the M–B chromaticity diagram. These dots are distributed approximately along the black body locus, shown by the red dashed line. Figure 2c shows the luminance distribution of the pixels on the redness $L/(L+M)$ and the blueness $S/(L+M)$ axis. It has the particular characteristics that pixels of higher luminance tend to have lower redness $L/(L+M)$ and higher blueness $S/(L+M)$, and that those of lower luminance tend to have higher redness and lower blueness. These two parts correspond with the cloth part (white/blue) and the lace part (gold/black) of the dress, respectively.

To apply the optimal color hypothesis, we sampled 20 points uniformly in space in each of the white or blue stripes and the gold or black stripes of the dress image. They are shown as blue circle and brown square symbols in Figures 2b and c. Two large cross symbols in Figures 2b and c represent two means of the sampled parts. For the sake of simplicity, we used these two means in our analysis to represent the white

and blue circles and crosses represent surface examples in a scene (see text for details). (c) 574 natural objects (Brown, 2003) and surfaces in 16 hyperspectral images (Foster et al., 2006) plotted in the optimal color shell of 6500K. Colors of dots for natural objects and surfaces roughly correspond with color-appearance of those regions in the M–B chromaticity diagram.

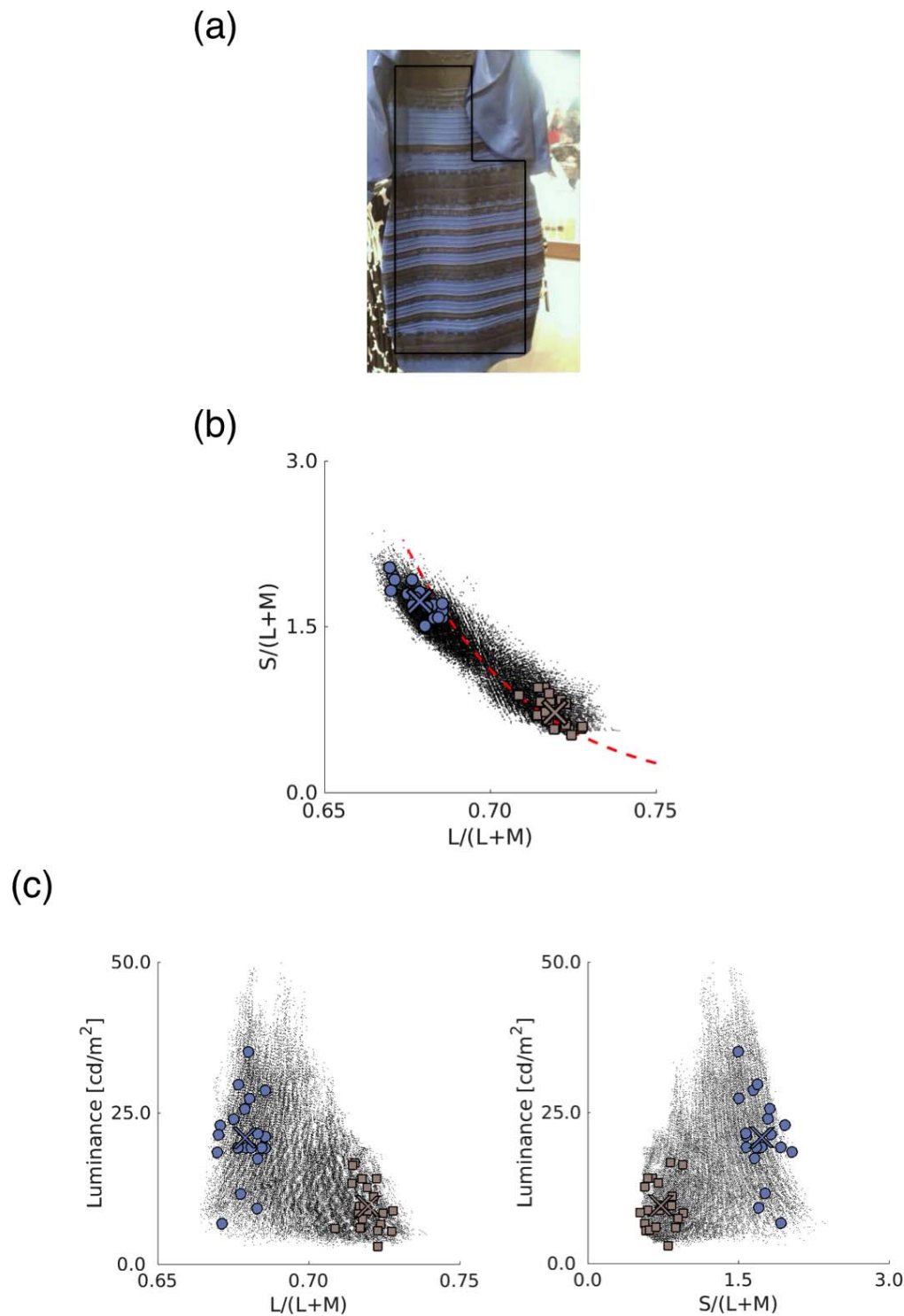


Figure 2. (a) #TheDress image shown with the sampling area surrounded by black lines. (b) Chromaticities of the sampled pixels in MacLeod–Boynton diagram (black dots). The red dashed line represents the black body locus. (c) Luminance distribution of the pixels on the redness $L / (L + M)$ and the blueness $S / (L + M)$ axis. Blue circle and brown square symbols show 20 points sampled uniformly in each of the white/blue part and the gold/black part of the dress image. Large cross symbols represent the means of each part's sampled points. Photograph of the dress used with permission. Copyright Cecilia Bleasdale.

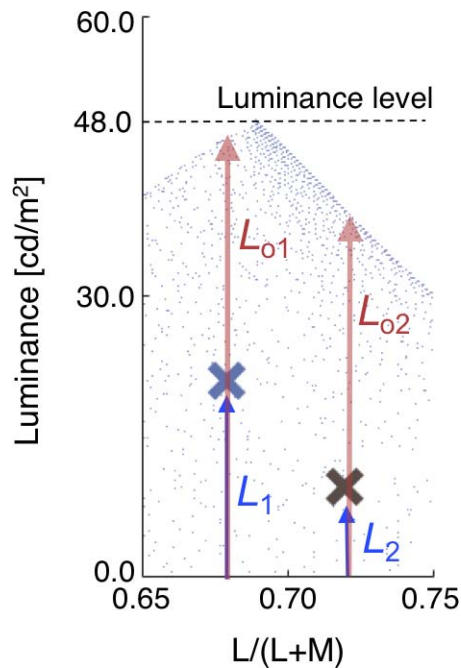


Figure 3. An example of an optimal color shell at a fixed luminance level of 48 cd/m^2 . Large cross symbols represent the means of each of the 20 points sampled uniformly in each of the white or blue part and the gold or black part of the dress image. L_i : luminance of one of the two means, L_{oi} : luminance of the optimal color with the same chromaticity, $i = 1, 2$.

or blue color and the gold or black color in the dress image.

When we determined the best-fit optimal color shell, we first fixed the luminance level for all optimal color shells with different color temperatures. Figure 3 shows an example where the luminance level is fixed at 48 cd/m^2 . Secondly, the luminance error e_i was calculated between the luminance of one of the two means L_i and luminance of the optimal color with the same chromaticity L_{oi} , where

L_i = luminance of one of the two means,

L_{oi} = luminance of the optimal color
with the same chromaticity,

$$e_i = (L_i - L_{oi})^2,$$

$i = 1, 2$.

The error was weighted by a factor f_i , which was used to enhance errors of higher luminance points (Uchikawa et al., 2012).

$$f_i = L_i / L_{oi}$$

Finally, the root-mean-square-error (RMSE) was

calculated for all optimal color shells of different color temperatures ranging from 3000K to 20000K in 500 K steps.

$$\text{RMSE} = \{(f_1 e_1 + f_2 e_2) / 2\}^{1/2}$$

Figure 4 shows the case of a low luminance level, 22 cd/m^2 . When the color temperature of the illuminant was low (3000K) one of the mean points (blue cross) exceeded the optimal color shell of 3000K, making this color temperature invalid since a stimulus outside of the optimal color shell would not appear as a surface, but as a luminosity (Fukuda, Numata, & Uchikawa, 2013; Speigle & Brainard, 1996; Uchikawa, Koida, Meguro, Yamauchi, & Kuriki, 2001). As the color temperature increased to 6500K at the same luminance level, the invalid mean point came closer to the optimal color shell of 6500K. When the color temperature becomes 20000K, the invalid mean point was now included inside of the optimal color shell of 20000K, making this color temperature valid. In the all valid color temperatures at this luminance level, the least RMSE was found at the color temperature of 20000K. For this case, the best-fit illuminant is 20000K.

The best-fit color temperature was obtained using the same procedure for other luminance levels. Figure 5 shows four examples of best-fit optimal color shells when the luminance level increased from 22 to 55 cd/m^2 , and RMSEs as functions of color temperature at four luminance levels. These optimal color shells gave the best-fit color temperatures.

Accordingly, the best-fit color temperature changed depending on the luminance level. In Figure 6, RMSEs were plotted as functions of color temperature at different luminance levels. The best-fit color temperature was found to be higher at a lower luminance level as shown by a blue arrow in the figure. When the luminance level increased by just a few cd/m^2 , from 22 to 26 cd/m^2 , the best-fit color temperature jumped to the lower value as shown by a white arrow. Then, the best-fit color temperature gradually decreased as the luminance level increased. Figure 6 shows that these functions have local maxima at around 13000K at luminance levels over 22 cd/m^2 , which is shown by a dotted black line. The shift due to illuminant intensity changes is a possible explanation of the observer-dependent estimation of the illuminant in the dress image.

When the illuminant was estimated to have low luminance and high color temperature, for example, in the case of 22 cd/m^2 and 20000K as seen in Figure 5, the dress should appear white/gold. This is because the mean chromaticity of the cloth part of the dress comes close to the peak of the optimal color shell, which corresponds to the white surface, and the mean chromaticity of the lace part is located in the orange or

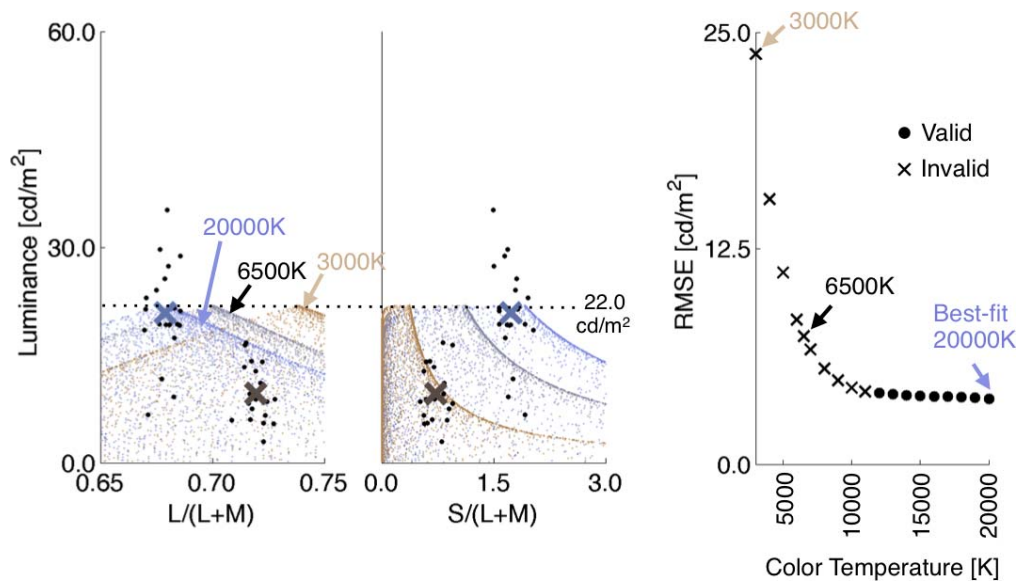


Figure 4. A graphical outline of obtaining the best-fit optimal color shell. The luminance level is fixed at 22 cd/m^2 . The optimal color shells of 3000K, 6500K, and 20000K are shown. Blue and brown crosses represent means of the white or blue and gold or black parts of the dress image, respectively. One of the mean points (blue cross) exceeds both optimal color shells of 3000K and 6500K, which makes it invalid. The right-most panel shows RMSE (root mean square error) for each color temperature. Closed circles and black crosses in the panel denote valid and invalid color temperatures, respectively.

brown region of the chromaticity diagram. On the other hand, when the illuminant was estimated to have high luminance and low color temperature, for example, in the case of 44 cd/m^2 and 5000K in Figure 5, the dress should appear blue/black, because the chromaticity of the cloth part is located in the dark blue region and that of the lace part localizes to the dark gray or black region compared with the peak of the 5000K optimal color shell.

We applied the optimal color hypothesis to the chromatically inverted dress to see whether its color appearance is also explained by this hypothesis. As shown in Figure 7a, the chromaticity of the dress was inverted by 180° on the redness and blueness axes of the M–B chromaticity diagram with no luminance change. The dress appears blue/yellow with almost no ambiguity. In Figure 7b, the blue and brown cross symbols represent means of 20 points (small dots) sampled in

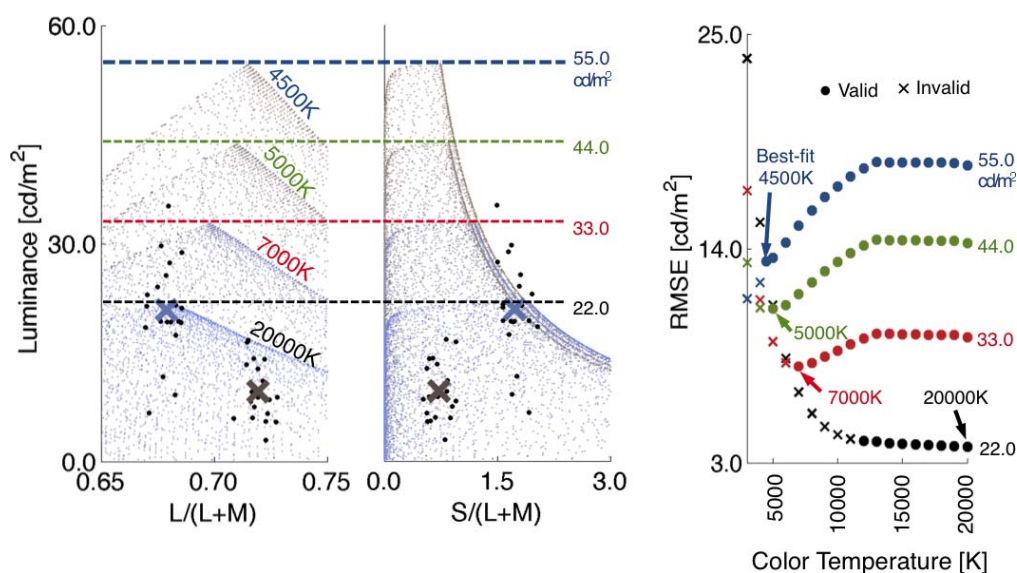


Figure 5. Four examples of best-fit optimal color shells and RMSEs as functions of color temperature when the luminance level increased from 22 to 55 cd/m^2 . Circles and crosses in the right-most panel represent valid and invalid color temperatures, respectively.

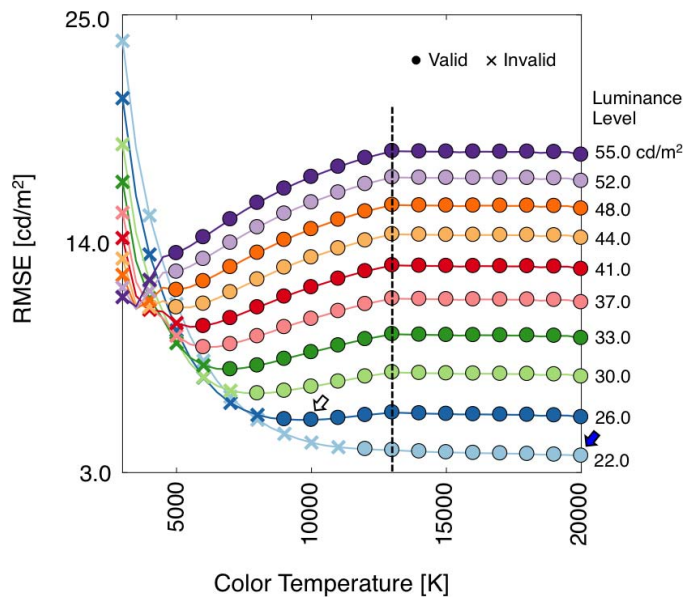


Figure 6. RMSEs plotted as functions of color temperature at different luminance levels. Circles and crosses represent valid and invalid color temperatures, respectively. The blue arrow shows the best-fit color temperature at 22 cd/m² and the white arrow shows the best-fit color temperature at 26 cd/m². The dotted black line shows local maxima at around 13000 K at luminance levels above 22 cd/m².

the lace part and that of the cloth part of the inverted dress image, respectively. The optimal color hypothesis predicted illuminants of low color temperatures between 4000 K and 3000 K no matter which luminance-level was used, as shown in Figure 7c. Under these illuminants, the dress should appear dark blue and bright yellow. This agrees well with the observed appearance of the dress.

In summary, in this analysis we found that the best-fit optimal color distribution depended on the illuminant luminance level. The best-fit illuminant color temperature changes abruptly from high to low at a specific luminance-level. From these findings, it might be the case that this luminance-dependent illuminant prediction may cause observer-dependent appearances of the dress.

Experiment

Method

Apparatus and stimuli

The stimulus was presented on a CRT monitor (Sony GDM-520, 19 in., 1024 × 768 pixels; Sony, Tokyo, Japan) controlled by a computer (Epson MT7500; Epson, Nagao, Japan) equipped with a

graphic card (ViSaGe, Cambridge Research System, Rochester, UK) of 14-bit intensity resolution for each phosphor. We calibrated each phosphor using the Color-CAL colorimeter (Cambridge Research System) for gamma correction and a spectral radiometer (PR-650, Photo Research Inc., New York, NY) for spectral energy distribution.

We used three dress images with different colors as stimuli. They were: original colors, chromatically inverted colors, and achromatic colors (as shown in Figure 8). The luminance distributions are shown in the right two panels for each dress image. The viewing distance was 114 cm. The square test field was made on the dress pattern with the luminance texture preserved as shown in the top panel of Figure 8. It subtended 2.0 × 3.6° of visual angle.

Procedure

The observer adjusted the chromaticity and luminance of the test field so that it appeared as a full-white surface. The observer's criterion was an achromatic paper match, where the test field appeared as if it was the brightest white surface (the white surface with maximum lightness in the surface-color mode) under this illuminant. The chromaticity and luminance of the test field set by the observer represents the chromaticity and luminance of an estimated illuminant.

In a block, the observer, first, named colors of the lace and cloth parts of a dress stimulus using Berlin and Kay's 11 basic color terms (white, black, red, green, yellow, blue, brown, orange, purple, pink, and gray) in addition to gold, silver, and copper. Then, the observer started adjusting the chromaticity and luminance of the test field so that it appeared as a full-white surface. This adjustment was repeated five times. In the next block, the dress stimulus was changed to one of the other dress stimuli selected at random. Three blocks were done for all three dress patterns in a session. The observer carried out four sessions in total in the experiment.

Observers

Two groups of five observers were used in the experiment. Observers were classified into two groups depending on whether they saw the original dress image either as white/gold (observers S1 to S5) or blue/black (observers S6 to S10). Their ages were 33, 23, 25, 38, and 35 (mean = 30.8) for observers S1, S2, S3, S4, and S5 in the white/gold group and 26, 33, 23, 27, and 18 (mean = 25.4) for observers S6, S7, S8, S9, and S10 in the blue/black group. The difference in age between the two groups was not significant ($p = 0.1939$). All observers were color normal as tested by the Ishihara color vision test.

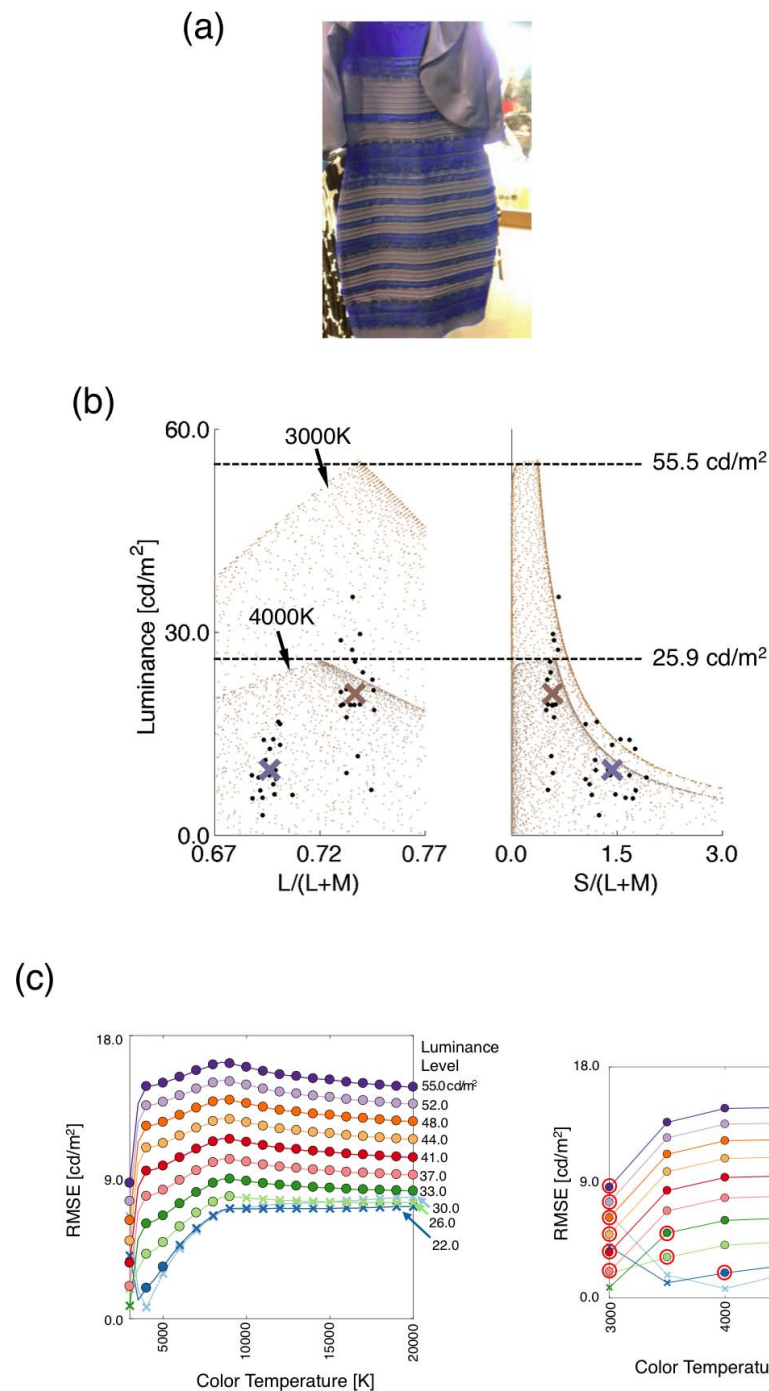


Figure 7. (a) The chromatically inverted dress image used in the analysis. (b) 20 points (small dots) sampled each in the lace part and the cloth part of the inverted dress image. Blue and brown cross symbols represent means of these 20 points. (c) RMSEs plotted as functions of color temperature at different luminance levels for the chromatically inverted dress. Circles and crosses represent valid and invalid color temperatures, respectively. In the right-hand panel the graph region at low color temperatures is enlarged. The thin line circles show minimum points at each luminance level. Photograph of the dress used with permission. Copyright Cecilia Bleasdale.

Results and discussion

Figures 9a, b, and c show the results for the original, inverted, and achromatic dress stimuli, respectively. Color names are shown in the top panel of each figure for all observers. A color name for the cloth part and

that for the lace part are separated by a slash in a parenthetical, for example, (white/gold) or (blue/black). It is shown in Figure 9a that five observers, S1 to S5, used white/gold (or orange for S2) and other five observers, S6 to S10, used blue/black for the original dress stimulus. Observers S1 to S5 were classified into

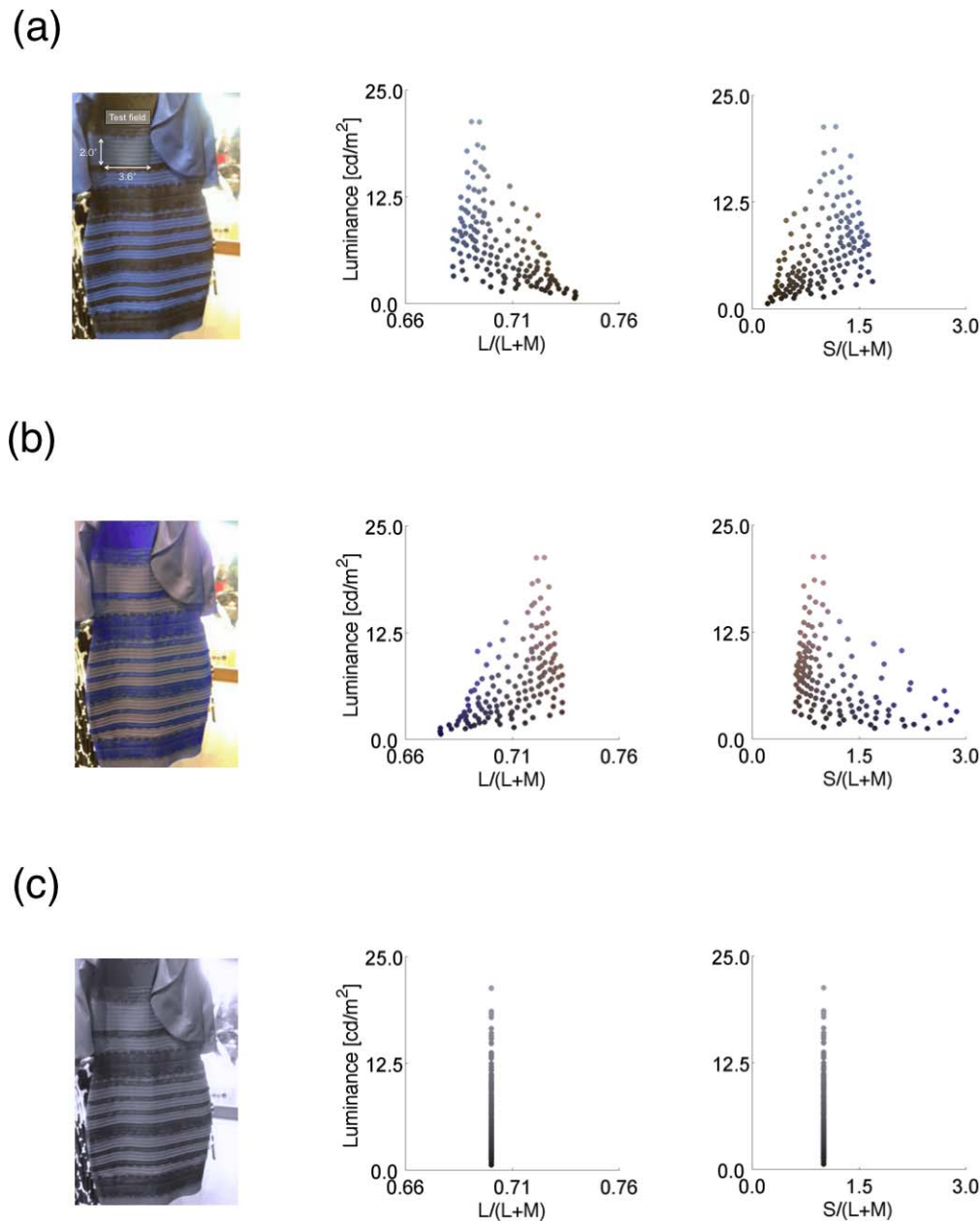


Figure 8. Three dress images with different colors used as stimuli. (a): original colors, (b): chromatically inverted colors, and (c): achromatic colors. The luminance distributions of pixels are shown in the right panel of each figure. Photograph of the dress used with permission. Copyright Cecilia Bleasdale.

the white/gold group and observers S6 to S10 into the blue/black group.

We plotted observers' chromaticity settings in a M–B chromaticity diagram as shown in lower left panels of Figures 9a, b, and c. The dotted line represents the black-body locus. Small black circles on the dotted line represent color temperatures in 500K steps. The plus symbol indicates the chromaticity of the equal energy white. Two cross symbols denote mean chromaticities across observers for each group. Observers' luminance setting, which was a mean luminance across all pixels within the test field, is shown in lower right panels of

Figures 9a, b, and c. Two dashed lines represent the mean luminances across observers for the two groups.

It is shown in Figure 9a that the white/gold group estimated the illuminant to be bluish, and the blue/black group estimated it to be yellowish for the original dress stimulus. The mean chromaticities for two groups were significantly different both for redness [$L / (L + M)$] and blueness [$S / (L + M)$] directions in the chromaticity diagram, $t(8) = 2.77$, $p = 0.024$, $d = 1.96$; $t(8) = 3.26$, $p = 0.012$, $d = 2.31$, respectively. In the luminance plot we can notice that luminance settings are quite well separated between the two groups. The

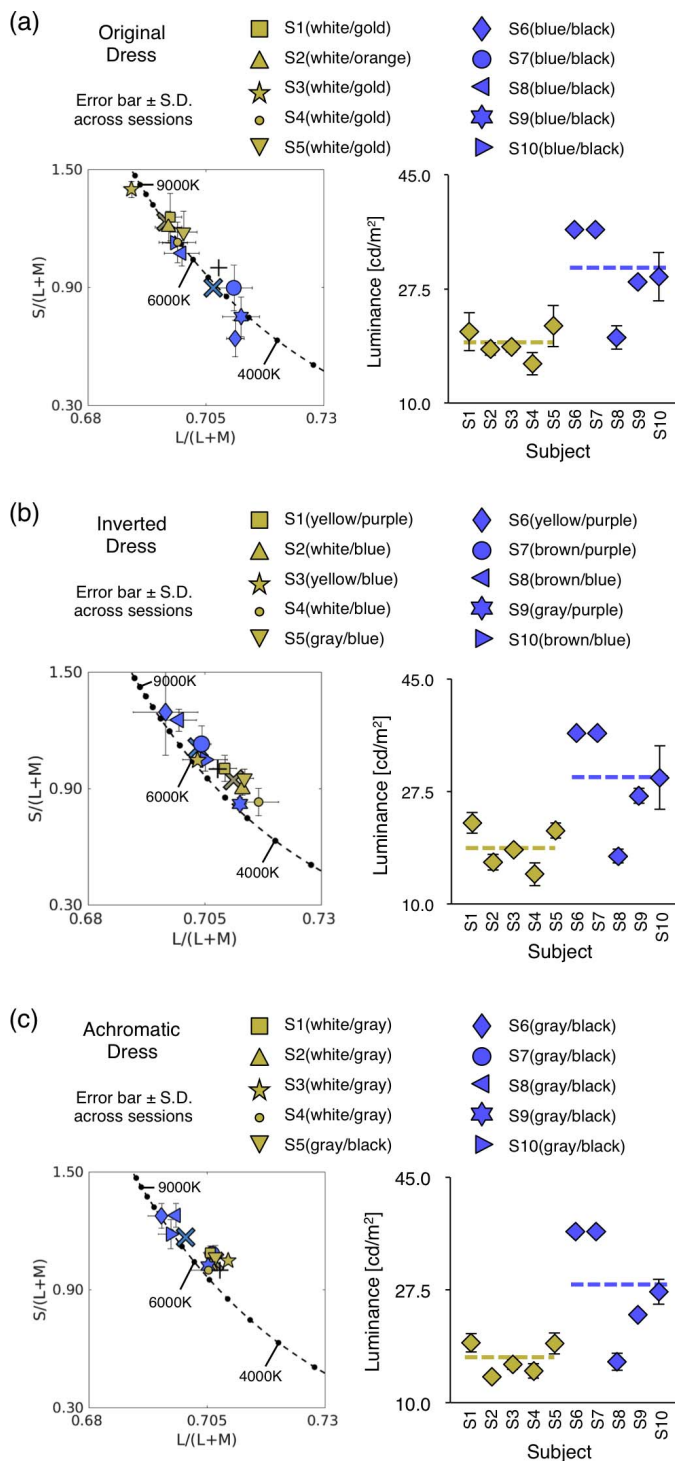


Figure 9. Color names used by each observer, observers' chromaticity settings in MacLeod–Boynton chromaticity diagram, and observers' luminance settings (averaged across the pixels within the test field) in the top, lower left and lower right panel for (a): the original dress, (b): chromatically inverted dress, and (c): achromatic dress stimuli. Subjects S1 to S5 designated with yellow symbols are in the white/gold group and observers S6 to S10 designated with blue symbols are in the blue/black group. Small black dots on the black-body locus shown as a dashed line corresponds with the color temperature from 3500K to 9500 K with 500 K steps.

difference in mean luminance was significant, $t(8) = 3.43$, $p = 0.0089$, $d = 2.43$. The white/gold group estimated the illuminant to be darker, and the blue/black group estimated it to be brighter. These estimates for chromaticity and luminance of the illuminant by the two groups are consistent with the predictions of optimal color theory.

Figure 9b shows the case of the chromatically inverted dress stimulus. The color names used by the observers were fairly consistent across all observers, that is, yellow, brown, white, or gray for the cloth part and purple or blue for the lace part regardless of the type of group. Observers' estimated chromaticities of the illuminant tend to shift toward the purple region (the upper-right region, in the M–B chromaticity diagram). They were not consistent, and the difference in the mean chromaticities for white/gold and blue/black groups was not significant neither in the redness, $t(8) = 2.15$, $p = 0.063$, nor blueness directions, $t(8) = 1.77$, $p = 0.12$. Interestingly, we can notice that observers' estimated luminances tended to be separated between the two groups in the luminance plot. This difference in luminance was found to be significant, $t(8) = 2.775$, $p = 0.024$, $d = 1.96$. They seem to appear darker for the white/gold group, and brighter for the blue/black group. This difference in luminance is quite similar to that estimated by the same observers for the original dress. When these results are compared with the predictions in Figure 7b we notice that chromaticities of the estimated illuminants vary more widely than the prediction. We will discuss this point later in the General discussion.

For the case of the achromatic dress stimulus shown in Figure 9c, all color names used in the white/gold group were white/gray except S5, but those used in the blue/black group were gray/black. Thus, the usage of color names clearly separated between the two groups. Chromaticities of the illuminant estimated by most observers were consistently located around the equal-energy white point except for three observers. The mean chromaticities for two groups were significantly different only in the redness direction, $t(8) = 2.68$, $p = 0.028$, $d = 1.89$, but not in the blueness direction, $t(8) = 2.29$, $p = 0.051$. It is, again, interesting that luminance settings were lower for the white/gold group, but they are higher for the blue/black group. This difference in the mean luminance was significant, $t(8) = 2.44$, $p = 0.041$, $d = 1.99$. This difference is quite similar to the case of the original dress, and very consistent with the usage of color names in two groups. It should be noted that in this condition the dress stimulus was achromatic, but the two groups estimated luminance of the illuminant to be different.

In summary, we found in this experiment that the white/gold group estimated the illuminant to be darker and bluish and the blue/black group estimated the

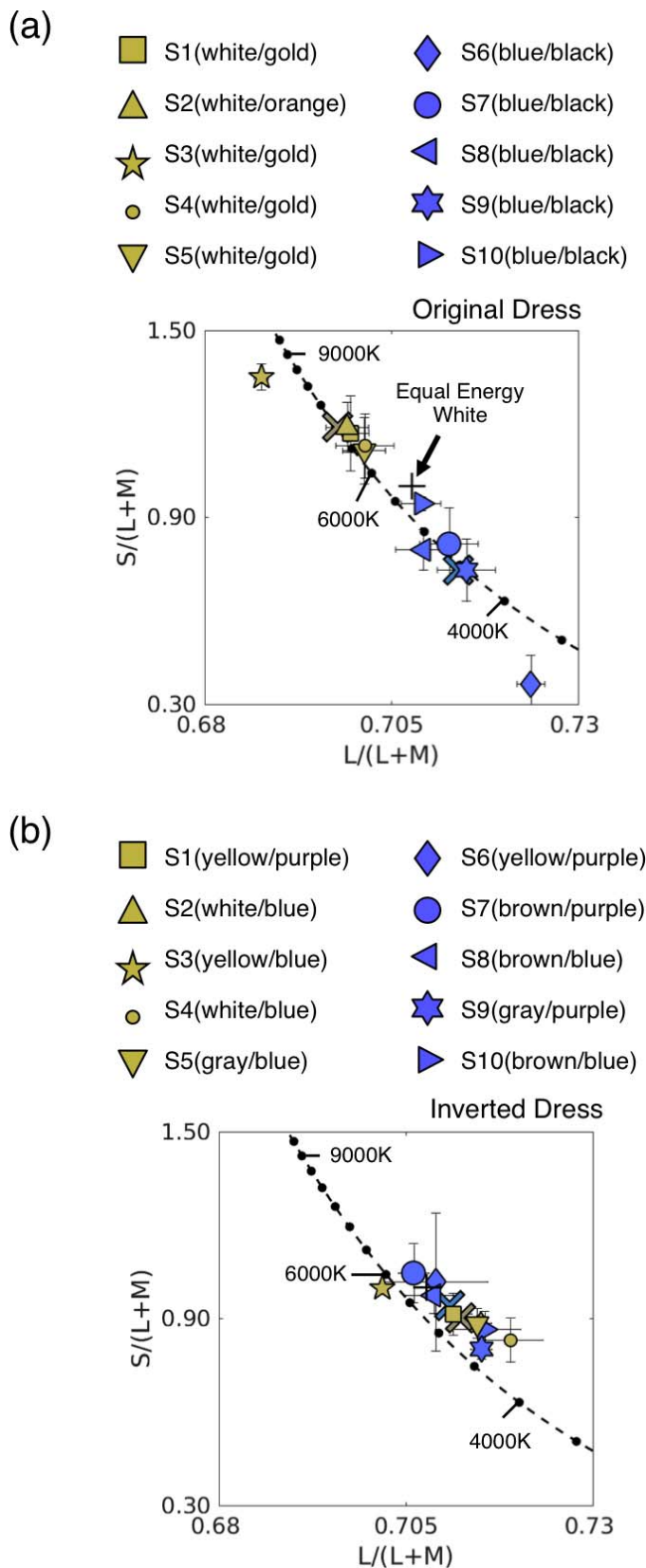


Figure 10. Observers' chromaticity settings shifted to cancel individual biases in MacLeod–Boynton chromaticity diagram. (a): the original dress and (b): chromatically inverted dress stimulus. Symbols correspond with observers in the same way

illuminant to be brighter and yellowish for the original dress. The observer's estimated illuminant was consistent with the predicted illuminant of optimal color theory. A new finding in this experiment was that even when the dress was achromatic, these two groups estimated the illuminant to be darker or brighter in the same way as they did for the original dress.

General discussion

It was shown that the white/gold and blue/black appearances of the dress could be explained by the luminance-dependent illuminant prediction by our optimal color hypothesis. The optimal color hypothesis predicts, in general, luminance as well as chromaticity of an illuminant in a scene (Morimoto, Fukuda, & Uchikawa, 2016; Uchikawa, Fukuda, Kitazawa, & MacLeod, 2012). In the present analysis, we first fixed the luminance level of an illuminant to see what chromaticity of the illuminant was chosen as the best chromaticity for the luminance level. Using this application of the optimal color hypothesis we could find the luminance-dependent illuminant characteristic of the dress image.

We found in the experiment that the white/gold and blue/black groups estimated luminance level of the illuminant to be darker and brighter, respectively, for even the “achromatic” dress in a similar way to the original dress. This finding might suggest that the visual system could first perceive the luminance level of an illuminant, and then determine the chromaticity of the illuminant in a scene. This suggestion would provide a rationale for searching for the chromaticity of the illuminant under a certain fixed luminance level in the optimal color hypothesis. It has been reported in other studies that assuming illumination intensity in the visual system was important to determine the appearance of the dress (Wallisch, 2017; Witzel, Racey, & O'Regan, 2017).

In Figures 9a, b, and c, we can see that an observer's settings of illuminant chromaticity spread along the black-body locus with wider variations for the blue/black group than for the white/gold group. In the white/gold group, the chromaticity points appeared to coincide with the predictions of the optimal color hypothesis both for the original and inverted dress images. However, the chromaticity points in the blue/black group appeared not to follow the predictions. For example, in Figure 9a, the chromaticity points of S8 and S10 were found to deviate from low color

as in Figure 9. Small black dots on the black-body locus shown as a dashed line corresponds to the color temperature from 3500K to 9500K with 500K steps.

temperature regions of the black-body locus, and instead were located in high color temperature regions close to the white/gold group. Additionally, in Figure 9b, the chromaticity points of S6 and S8 also deviated from other observers' points to higher color temperature regions.

It is likely that observers have their own biases of white points in a chromaticity diagram. Bosten, Beer, and MacLeod (2015) reported a fairly large deviation of individual settings of white ranging approximately from 5000K to 8000K along the daylight locus. This deviation corresponds well with those observed in Figures 9a, b, and c. The variation in our data might not be due to prediction failure by the optimal color hypothesis, but due to the individual bias of an observer's white point. In fact, no external reference for our experimental task was given to observers.

To try to cancel the individual bias of white judgment of an illuminant we used the chromaticity points obtained for the achromatic dress image in Figure 9c as reference of an individual's bias of a white surface. In Figure 9c, we calculated the Euclidean distance between each observer's setting point and the equal-energy white point, and also recorded the direction of each data point from the equal-energy white point in the M–B chromaticity diagram. Then all data points in Figures 9a and b were shifted in the opposite direction with the same distance in the M–B chromaticity diagram. Figures 10a and b shows the chromaticities of observers' settings shifted to cancel individual biases of white. In Figure 10a the data points appeared to be separated quite clearly between the two groups, which would better support the predictions of the optimal color hypothesis. In Figure 10b, all data points in either group came to the same narrow region between 4500K and 6000K, which would again better support our predictions.

Our optimal color hypothesis infers the most suitable illuminant in a scene from only the luminance versus chromaticity distribution in the scene. The dress image consists of two color regions (bright blue and dark yellow) which create an unnatural distribution of luminance (Figure 2); whereas a natural scene in our environments tends to contain dark blue and bright yellow surfaces as shown in Figure 1c. We conclude that this unnatural distribution of luminance in the dress image is likely one of the factors, which made the dress phenomenon so special because the visual system might hardly find one stable solution for illuminant estimation in the dress image.

However, the luminance versus chromaticity distribution alone may not explain the whole story in the case of the dress phenomenon. In the present experiment, the observers estimated the luminance level of the illuminant to be bright or dark even for the achromatic dress almost in the same way as for the original dress

(Figure 9c). This would imply that the luminance level of the illuminant must be determined differently for individuals using other factors.

It would be plausible to assume that the other factors might be some luminance cues hidden in the dress image because in our experiment, illuminant luminance levels estimated by each observer were almost constant for the original, inverted, and achromatic dress images. In these dress images the luminance pattern of the dress images was kept the same.

These cues might work differently in each observer on the basis of their environmental experiences as suggested in Mahroo et al. (2017). They reported in their twin study that environmental factors played a significant role as opposed to genetic factors to determine individual color appearance of the dress. The question of what such environmental factors might be is still unresolved.

We can easily recognize the lace and cloth parts as surfaces in the original, inverted, and achromatic dress images. Hesslinger and Carbon (2016) showed that when the spatial structure of the dress image was broken down by means of image scrambling the perceptual difference between white/gold and blue/black observers decreased significantly. These results indicate that the spatial structure of the image was needed for the visual system to see surfaces in the image and obtain illumination information about those the surfaces. Matsumoto, Morimoto, and Uchikawa (2016) reported that when a striped part of the dress pattern ($9.1^\circ \times 5.4^\circ$), cut from the original image ($15.2^\circ \times 23.0^\circ$), was used as a stimulus, the illuminant estimations of white/gold and blue/black groups were not clearly separated. When a spatially uniform, textureless, stripe pattern with the same chromaticities was used, the separation between the two groups was further degraded. These results suggest that some spatial structures of the dress image are crucial factors in the dress phenomenon.

Keywords: dress, color constancy, optimal color hypothesis, illuminant estimation

Acknowledgments

This research was supported by JSPS KAKENHI 25245065 and 15K12125. The authors are grateful to Kenji Kagitani for his comments on “#TheDress,” and to Tanner DeLawyer for his careful correction of grammatical errors in the text of the manuscript.

Commercial relationships: none.

Corresponding author: Keiji Uchikawa.

Email: uchikawa@ip.titech.ac.jp.

Address: Research Institute for Multimodal Sensory Science, Faculty of Human Sciences, Kanagawa University, Yokohama, Japan.

References

- Bosten, J. M., Beer, R. D., & MacLeod, D. I. A. (2015). What is white? *Journal of Vision*, 15(16):5, 1–19, doi:10.1167/15.16.5. [PubMed] [Article]
- Brainard, D. H., & Hurlbert, A. C. (2015). Colour Vision: Understanding #TheDress. *Current Biology*, 25(13), 551–554, doi:10.1016/j.cub.2015.05.020.
- Brown, R. (2003). Background and illuminants: The yin and yang of colour constancy. In R. Mausfeld & D. Heyer (Eds.), *Colour perception: Mind and the physical world* (pp. 247–272). New York: Oxford University Press.
- Churma, M. E. (1994). Blue shadows: Physical, physiological, and psychological causes. *Applied Optics*, 33(21), 4719–4722, doi:10.1364/AO.33.004719.
- Foster, D. H., Amano, K., Nascimento, S. M. C., & Foster, M. J. (2006). Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A*, 23(10), 2359–2372, doi:10.1364/JOSAA.23.002359.
- Foster, D. H. (2011). Color constancy. *Vision Research*, 51, 674–700, doi:10.1016/j.visres.2010.09.006
- Fukuda, K., Numata, A. & Uchikawa, K. (2013). Perceptual luminosity threshold on the surround stimulus that simulates luminance-chromaticity distributions in natural scene. *Journal of Vision*, 13(9): 1020, doi:10.1167/13.9.1020. [Abstract]
- Fukuda, K., & Uchikawa, K. (2014). Color constancy in a scene with bright colors that do not have a fully natural surface appearance. *Journal of the Optical Society of America A*, 31(4), 239–246, doi:10.1364/JOSAA.31.00A239.
- Gegenfurtner, K. R., Bloj, M., & Toscani, M. (2015). The many colours of ‘the dress.’ *Current Biology*, 25(13), 543–544, doi:10.1016/j.cub.2015.04.043.
- Hesslinger, V. M., & Carbon, C. C. (2016). #TheDress: The role of illumination information and individual differences in the psychophysics of perceiving white-blue ambiguities. *i-Perception*, 7(2), 1–10, doi:10.1177/2041669516645592.
- Hurlbert, A. (2007). Colour constancy. *Current Biology*, 17(21), 906–907, doi:10.1016/j.cub.2007.08.022.
- Lafer-Sousa, R., Hermann, K. L., & Conway, B. R. (2015). Striking individual differences in color perception uncovered by ‘the dress’ photograph. *Current Biology*, 25(13), 545–546, doi:10.1016/j.cub.2015.04.053.
- Mahroo, O. A., Williams, K. M., Hossain, I. T., Yonova-Doing, E., Kozareva, D., Yusuf, A., ... Hammond, C. J. (2017). Do twins share the same dress code? Quantifying relative genetic and environmental contributions to subjective perceptions of “the dress” in a classical twin study. *Journal of Vision*, 17(1):29, 1–7, doi:10.1167/17.1.29. [PubMed] [Article]
- Matsumoto T., Morimoto T., & Uchikawa K. (2016). Estimation of illuminant in the white/gold and blue/black “dress” image. *Journal of Vision*, 16(12): 220, doi:10.1167/16.12.220. [Abstract]
- Melgosa, M., Gómez-Robledo, L., Suero, M. I., & Fairchild, M. D. (2015). What can we learn from a dress with ambiguous colors? *Color Research & Application*, 40(5), 525–529, doi:10.1002/col.21966.
- Morimoto, T., Fukuda, K., & Uchikawa, K. (2016). Effects of surrounding stimulus properties on color constancy based on luminance balance. *Journal of the Optical Society of America A*, 33(3), 214–227, doi:10.1364/JOSAA.33.00A214.
- Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G. D., & Hurlbert, A. (2014). Chromatic illumination discrimination ability reveals that human colour constancy is optimized for blue daylight illuminations. *PLoS One*, 9(2), e87989, doi:10.1371/journal.pone.0087989.
- Schlaefke, L., Golisch, A., Haag, L. M., Lenz, M., Heba, S., Lissek, S., ... Tegenthoff, M. (2015). The brain’s dress code: How The Dress allows to decode the neuronal pathway of an optical illusion. *Cortex*, 73, 271–275, doi:10.1016/j.cortex.2015.08.017.
- Speigle, J. M., & Brainard, D. H. (1996). Luminosity thresholds: Effects of test chromaticity and ambient illumination. *Journal of the Optical Society of America A*, 13, 436–451.
- Stockman, A., & Sharpe, L.T. (2000). Tritanopic color matches and the middle- and long-wavelength-sensitive cone spectral sensitivities. *Vision Research*, 40, 1739–1750.
- Uchikawa, K., Koida, K., Meguro, T., Yamauchi, Y., & Kuriki, I. (2001). Brightness, not luminance, determines transition from the surface-color to the aperture-color mode for colored lights. *Journal of the Optical Society of America A*, 18(4), 737–746.
- Uchikawa, K., Fukuda, K., Kitazawa, Y., & MacLeod, D. I. A. (2012). Estimating illuminant color based on luminance balance of surfaces. *Journal of Optical Society of America A*, 29(2), 133–143, doi:10.1364/JOSAA.29.00A133.
- Vemuri, K., Bisla, K., Mulpuru, S., & Varadharajan,

- S. (2016). Do normal pupil diameter differences in the population underlie the color selection of #thedress? *Journal of the Optical Society of America A*, 33(3), 137–142, doi:10.1364/JOSAA.33.00A137.
- Wallisch, P. (2017). Illumination assumptions account for individual differences in the perceptual interpretation of a profoundly ambiguous stimulus in the color domain: “The dress”. *Journal of Vision*, 17(4):5, 1–14, doi:10.1167/17.4.5 [PubMed] [Article]
- Winkler, A., Spillmann, L., Werner, J.S., & Webster, M.A. (2015). Asymmetries in blue-yellow color perception and in the color of ‘the dress.’ *Current Biology*, 25, 1–2, doi:10.1016/j.cub.2015.05.004.
- Witzel, C., Racey, C., & O’Regan, J. K. (2017). The most reasonable explanation of “the dress”: Implicit assumptions about illumination. *Journal of Vision*, 17(2):1, 1–19, doi:10.1167/17.2.1. [PubMed] [Article]