Influence of Texture Structure on the Perception of Color Composition

Jing Wang[1], Jana Zujovic[2], June Choi[3], Basabdutta Chakraborty[4], Rene van Egmond[5], Huib de Ridder[5], Thrasyvoulos N. Pappas[6],▲

[1]*Bloomberg LP, New York, NY, USA;* [2]*Google, Mountain View, CA, USA;* [3]*Accenture, New York, NY, USA;* [4]*Amway, Ada, MI, USA;* [5]*Delft University of Technology, Delft, Netherlands;* [6]*Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208, USA*
E-mail: pappas@eecs.northwestern.edu

Abstract. *The authors explore the influence of the structure of a texture image on the perception of its color composition through a series of psychophysical studies. They estimate the color composition of a texture by extracting its dominant colors and the associated percentages. They then synthesize new textures with the same color composition but different geometric structural patterns. They conduct empirical studies in the form of two-alternative forced choice tests to determine the influence of two structural factors, pattern scale and shape, on the perceived amount of target color. The results of their studies indicate that (a) participants are able to consistently assess differences in color composition for textures of similar shape and scale, and (b) the perception of color composition is nonveridical. Pattern scale and shape have a strong influence on perceived color composition: the larger the scale, the higher the perceived amount of the target color, and the more elongated the shape, the lower the perceived amount of the target color. The authors also present a simple model that is consistent with the results of their empirical studies by accounting for the reduced visibility of the pixels near the color boundaries. In addition to a better understanding of human perception of color composition, their findings will contribute to the development of color texture similarity metrics.* Ⓒ *2020 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.Percept.Imaging.2020.3.1.010401]

1. INTRODUCTION

Texture is an important visual attribute that provides important cues for object boundary detection and localization, foreground/background separation, and material identification [3, 26]. In this article, we explore some aspects of texture perception that are important for the development of texture similarity metrics [45, 59, 60], which play a key role in engineering applications, such as image compression, restoration, content-based retrieval, and understanding [10, 30, 40, 45, 61]. In particular, we conduct a series of

psychophysical studies to examine the influence of the structure of a texture image on the perception of its color composition. By *color composition* we mean the amount (area in pixels) of each color in the image irrespective of position, and by *structure* we mean their spatial arrangement.

There is a large body of work on texture perception in the psychophysical literature. To a large extent, it has relied on computer-generated or carefully constructed images, from the pioneering work of Julesz [19, 21] and Beck [2] on pre-attentive pattern discrimination, to more recent studies of surface roughness perception [16, 18, 42] and surface gloss [1, 17, 34, 39, 58]. The use of synthetic images makes it possible to study specific aspects of texture and to isolate the parameters and neural mechanisms that affect its appearance. However, texture images that are encountered in real-world applications are more complex, involve several parameters, and as we will see, the study of their perceptual properties raises some new questions that have not been studied in the psychophysical literature.

1.1 Background

In the development of perceptual metrics for texture similarity, Zujovic et al. [59, 60] combined separate estimates of texture similarity in terms of structure and color composition. Taking into consideration the fact that textures can have similar structure and different color composition, as shown in Figure 1(a), or similar color composition and different structure, as shown in Fig. 1(b), they argued that how these separate estimates should be combined should depend on the observer and the application [59, 60].

Zujovic et al. [59, 60] argued that the structure of a texture can be reasonably approximated with the structure of the grayscale component of the image. While the chrominance also contributes to the structure, the case where structure is solely determined by the chrominance is possible [56] but unlikely in natural or synthetic textures. To evaluate the similarity of the structure of grayscale textures, they then
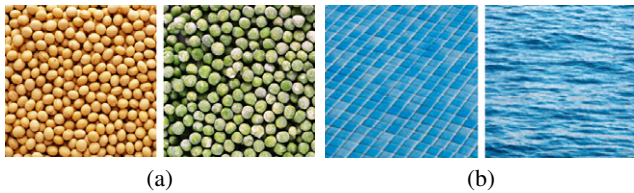

Figure 1. Examples of color textures with (a) similar structure and different color composition and (b) similar color composition and different structure.

used perceptually based *structural texture similarity metrics (STSIMs)* [61].

To estimate the color composition of a texture, as perceived by a human, rather than as a histogram of the colors of the pixels, Zujovic et al. [59, 60] adopted the representation in terms of dominant colors and their percentages, and used the *optimal color composition distance (OCCD)* [37], which is closely related to the earth mover's distance (EMD) [49], to compare the color compositions of two textures. Figure 2(a) shows an original texture image, Fig. 2(b) shows its grayscale component, and Fig. 2(c) shows its dominant colors. The representation of the color composition of a texture in terms of its dominant colors has been well established in the engineering literature [7, 29, 38, 59, 60]. We could not find any related work in the psychophysical literature, except for the work of Kuriki [25] on average-color perception of multicolored patterns. Kimura [22] also considered color averaging of multicolor mosaics into a few color categories, with emphasis on the mean color judgments.

1.2 *Proposed Approach*
The key assumption in Zujovic's approach to texture similarity is that texture structure and color composition are independent of each other. The goal of this article is to test whether this assumption is consistent with human perception and, in particular, whether the perception of color composition is veridical and whether it is affected by the texture structure.

We explore the influence of texture structure on the perception of the color composition through a series of empirical studies. While the motivation comes from texture similarity and content-based retrieval, the problem is interesting in its own right, and to the best of our knowledge has not been addressed in the psychophysical literature. Thus, while we will discuss the implications of the conclusions of our studies for texture similarity metrics, the main focus is on understanding the perception of the color composition of visual textures.

To study the effects of texture structure on color composition, we synthesized color textures with the same color composition as that extracted from the original texture but with different structural patterns, and conducted empirical studies to compare the perceived color composition of the synthesized textures.

As in [59, 60], we estimated the color composition of a given original texture image using the adaptive clustering algorithm (ACA) [43] to segment the image in $K_{ACA} = 4$

classes based on color, and "painting" the segments with the average of each class. ACA is an iterative algorithm that can be regarded as a generalization of the K-means clustering algorithm [28, 55] in two respects: it adapts to local variations in image intensity and includes spatial constraints in the form of Markov random fields (MRFs). We will refer to the resulting image, which preserves the structure of the original image, and consolidates the point clouds of each dominant color into points in color space, as the *posterized texture*. The posterized texture for the original texture image of Fig. 2(a) is shown in Fig. 2(c). The color composition of the texture consists of the (dominant) colors of this posterized texture and their percentages, shown in Fig. 2(d). We then synthesized textures with the same or modified color composition. We used three types of structural patterns for the synthetic textures: isotropic blobs, squares, and rectangles. As in the Julesz experiments [19], the synthesized textures were designed to eliminate familiarity cues, and the pattern shapes were chosen to create different structures (shape, scale) so that we can test their effect on the perceived color composition. An example of a synthetic texture with isotropic blobs is shown in Fig. 2(e).

We conducted three empirical studies with the synthetic textures to determine the effects of structure on the perception of color composition. In the first study, we compared the structure of the posterized texture with isotropic blob synthetic textures that have varying target color percentages. In the other two studies, we used square and rectangular synthetic textures to investigate the effect of scale and shape on the perceived amount of the target color. We conducted the empirical studies in the form of two-alternative forced choice (2AFC) tests to determine which of a pair of patterns is perceived as containing a larger amount of a given target dominant color. The main conclusions of our empirical studies are that (a) participants are able to consistently assess differences in color composition for textures of similar scale and shape, and (b) pattern agglomeration in both scale (larger) and shape (less elongated) has a strong positive influence on the perceived amount of a target color. Thus, the perception of color composition is nonveridical.

The question is then: What are the perceptual mechanisms that can account for the results of our empirical studies? First, the results are consistent with Julesz's texton theory [20], whereby the size and shape of the textons affect the texture perception. In Section 5, we present a simple model that estimates the contribution of each pixel to the perceived amount of the target color. Pixels near the blob boundaries contribute less than pixels at the center of the blobs. While there is a large literature on texture and color perception, we could not find any perceptual mechanisms that can explain such an edge effect. Here we should clarify that the size of the color blobs in our studies was selected so that the blobs are perceived as distinct patches of color with sharp edges. There is no averaging across the blob boundaries, even though a simple linear filter model based on estimates of the spatial frequency sensitivity of the eye (e.g., in [32]) predicts averaging over a few pixels at the given
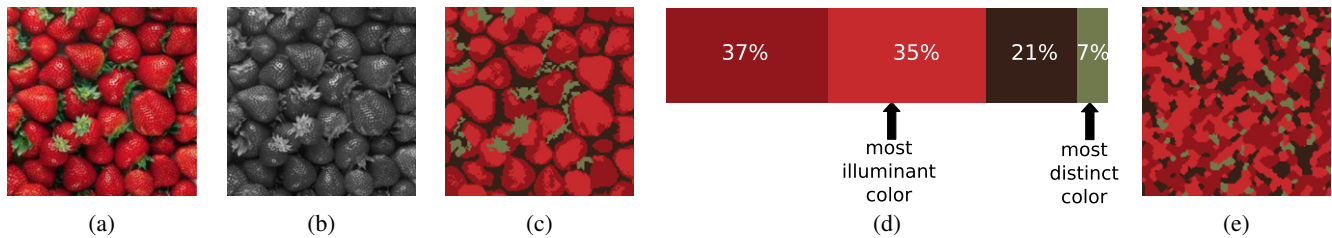
Figure 2. Texture structure and color composition. (a) Original image. (b) Grayscale texture of (a). (c) Posterized (texture segmentation with dominant colors). (d) Color composition, most illuminant, and most distinct color. (e) Synthetic texture based on MRF model with the same color composition as (c).

display resolution and viewing distance. The perception of sharp boundaries can be attributed to adaptation [12] or color spreading [47], but neither provides a solid explanation. In addition, with only a couple of exceptions, the differences in the colors of the blobs are well above threshold in both luminance and chrominance, so that differences in contrast sensitivity of luminance and chrominance edges [24] and changes in contrast sensitivity based on spatial pattern [48] or interactions between chrominance and luminance [53] cannot account for the observed results, and the effects of scale in particular. Another interesting observation is that, as in the case of texture metamers [8], all pixels in the color blobs are clearly visible, yet their contributions to the overall perception of color composition are different.

Webster et al. [57] and Maule and Franklin [36] conducted experiments with ensembles of distinct color blobs, like ours, but in regular formations of color circles. However, the emphasis was on the perception of the average color rather than the color composition.

1.3 Related Work

Objective descriptors of the color composition of an image have been extensively studied in the image retrieval literature over the past decades. The most straightforward way for describing the color content of an image is via a color histogram; two images are then compared using a histogram distance metric [50, 52]. Manjunath et al. [31] review three descriptors based on histogram representation for the MPEG-7 standard: the scalable color descriptor, the dominant color descriptor, and the color layout descriptor. The scalable color descriptor is a color histogram of an image encoded based on the Haar transform. The dominant color descriptor describes the colors of an image using a small number of dominant color values and the related statistical properties. The color layout descriptor captures the spatial distribution of dominant colors based on the discrete cosine transform (DCT).

The use of dominant colors and associated percentages as a compact color representation for image analysis was introduced by Ma et al. [29] and Deng et al. [9] and adopted by Mojsilovic et al. [38]. They argued that, when evaluating the color composition of an image, the visual system discounts local subtle variations and focuses on a few dominant colors.

In addition to an appropriate color representation, the color composition of images must be compared in a way that agrees with human perception. Mojsilovic et al. [37] proposed the OCCD, which is implemented in *CIELab* space and which is closely related to the EMD [49]. In the paper that introduced EMD, Rubner et al. [49] also emphasized the need for compact color representations, which is consistent with the use of dominant colors.

In line with these findings, and in order to account for the nonuniformity of the statistical characteristics of natural textures, Chen et al. [7] introduced the idea of spatially adaptive dominant colors in the context of color texture segmentation and proposed the use of ACA [43] to estimate them and OCCD [37] to compare them. Zujovic et al. [59, 60] then incorporated the spatially adaptive dominant colors and the OCCD into a structural texture similarity metric that, as we discussed above, assumes that color composition and texture structure can be estimated independently.

2. TEXTURE ANALYSIS AND SYNTHESIS

To analyze the effects of structure on the perceived color composition of a texture image, we first estimate its color composition in terms of dominant colors [7] and then synthesize textures with the same color composition and different structures.

2.1 Color Composition Feature Extraction

As we discussed in the introduction, we estimate the color composition of a texture image by extracting the dominant colors and the associated percentages, using ACA [43] to segment the image into regions of slowly varying colors with rapid changes at the boundaries. As we saw, this results in the posterized images of Fig. 2(c).

For spatially homogeneous textures, a small number of segment classes, typically $K_{ACA} = 2$ to 4, is sufficient to capture the dominant structure and colors of the image. Pappas et al. [44] found that, for natural images, the majority of segments containing perceptually uniform textures can be characterized by just the first two dominant colors for effective texture classification. Similarly, He and Pappas [14, 15] proposed a segmentation algorithm that is based on the fact that natural textures consist of intensity variations of a single hue [41]. However, as we discuss in Section 3, for our experiments we selected stimuli that have a variety of patterns and colors, with $K_{ACA} = 4$.

The feature vector that specifies the color composition, shown in Fig. 2(d), consists of the K_{ACA} averages, expressed

in CIELab color coordinates $\{L_k^*, a_k^*, b_k^*\}$ and the associated color percentages $\{p_k\}$ with $k = 1, \ldots, K_{ACA}$.

Among the extracted dominant colors, we define the *most illuminant* color as the color with the highest value in the luminance (L^*) channel and with percentage at least 20%. We also define the *most distinct* color as the color with the largest average ΔE distance from the other dominant colors. The ΔE distance between two colors (L_1^*, a_1^*, b_1^*) and (L_2^*, a_2^*, b_2^*) is determined as follows:

$$\Delta E_{Lab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}, \qquad (1)$$

where $\Delta L^* = L_1^* - L_2^*$, $\Delta a^* = a_1^* - a_2^*$, and $\Delta b^* = b_1^* - b_2^*$. The selection of the most illuminant and most distinct colors for the empirical studies we describe below was arbitrary. Initially, we thought that it would make it easier to run our studies. However, we have no evidence of that, and expect that similar outcomes would be observed with any other color choice.

2.2 Texture Generation

Our goal is to synthesize textures with a given color composition but with structure different from that of the posterized texture. We used two approaches for generating textures, one with isotropic blobs and the other with blocks of different geometries (squares, rectangles, lines) and scales.

2.2.1 Isotropic Blobs

To obtain textures that consist of blobs with colors and percentages that correspond to the given color composition, we rely on MRFs [4, 5, 23]. For more details, see appendix A.

For the first empirical study, we generated a number of blobby textures by varying the percentage of the target color, and adjusting the percentages of the other colors accordingly. For each set of colors and each set of percentages, we synthesized a new texture from scratch. Examples are shown in Figure 3(a).

2.2.2 Geometric Blocks

As an alternative to isotropic blobs, we consider the problem of synthesizing textures that consist of blocks with different geometric shapes and different scales. In principle, this could be done by adjusting the parameters of the MRF model; however, this is computationally intensive, and it is difficult to control the shape of the blocks. So, we adopt a more direct approach that places blocks of different colors and a given (rectangular) shape at different locations in the image; this is a manifestation of the "dead leaves" technique [6, 27, 35], and is designed to generate textures with a given scale and shape and a given color composition.

Given a set of colors and associated percentages, it is in general impossible to generate perfectly tiled images consisting of rectangular blocks of a fixed shape and size. Instead of perfectly tiled images, the idea is to place fixed shape/size blocks of different (dominant) colors at random positions in the image. This means, of course, that the blocks will overlap. To achieve the desired color percentages, we
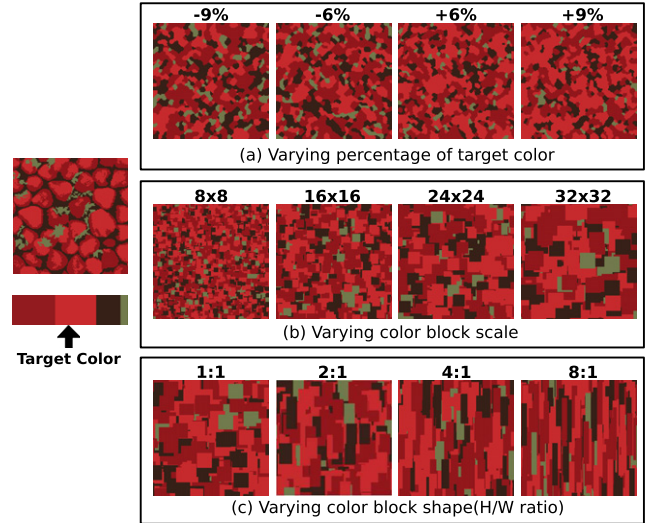

Figure 3. Synthesized images with (a) isotropic blobs with varying target color percentage, (b) squares of varying scale, and (c) rectangles with varying height/width ratio.

select the color of each block we place with probability equal to the percentage of that color. However, due to the random block overlap, the resulting percentages of each color will not be the same as the target percentages. We thus need an iterative procedure that will closely approximate the desired color percentages. The details can be found in appendix A. By selecting square blocks of a different fixed size for each image, we can get textures of different scales, as shown in Fig. 3(b). By controlling the height/width (H/W) ratio of the blocks, we can generate more or less elongated textures, as shown in Fig. 3(c).

When the percentage changes, the number of blocks of each color changes but, in principle, the shape and scale remain the same. However, due to the block overlap and random placement, this approach provides only approximate control of the scale and shape of the texture patterns. Quantitative analysis and illustrations of such effects can be found in Figure 11 of Study 2 and Figure 14 of Study 3.

An alternative approach to achieve a given color composition would be to keep the probability of placement fixed and to vary the size of the blocks based on the percentages. However, this approach does not provide good control of the scale of the resulting textures.

3. EMPIRICAL STUDIES

In this section, we describe three empirical studies to determine the effects of texture structure on the perception of color composition. In the first study, we compare the posterized texture with isotropic blob synthetic textures that have varying target color percentages, in order to determine whether there is a difference in the perceived target color percentages of the two texture structures. In the other two studies, we investigate the effect of scale and shape on the perception of the percentage of the target color. All three studies were designed as 2AFC tests.

3.1 *Test Setup*

3.1.1 *Participants*

Fifteen volunteers, six female and nine male, with ages ranging from 20 to 50, participated in each of the three studies. All participants had normal or corrected-to-normal vision and were tested for color blindness and red–green color vision deficiency. Before the test, all participants were asked to read and sign a consent form.

3.1.2 *Apparatus*

The tests were conducted using a calibrated LCD screen with 1920×1080 resolution and linear gamma. The viewing distance was approximately 60 cm so that a 256 pixel wide image subtended an angle of 9.39 degrees.

3.1.3 *Texture Stimuli*

For our empirical studies, we selected seven original full color textures, shown in Figure 4(a). The textures were obtained from the Corbis database,[1] correspond to real-world images, and were selected to have a variety of interesting patterns and colors. Since most natural textures have only a couple of dominant colors [15, 44] that typically consist of intensity variations (caused by changes in illumination) of a single hue (which corresponds to a single material) [41], we selected textures of flowers, fruits, corals, and fabrics that have four distinct dominant colors, as well as a variety of patterns. The posterized textures, obtained by ACA segmentation of the original textures, are shown in Fig. 4(b). Fig. 4(c) shows the color composition of the posterized textures of Fig. 4(b) with the target color underlined. (Note that Texture 5 contains a small amount of a fourth color, gray, that is visible at the upper left and right corners.) The length of each color bar represents the percentage. Figure 5(d) shows the grayscale component (loosely, luminance) of the posterized textures. Observe that the dominant colors of each texture differ in luminance, with two exceptions: Two of the colors in Texture Sets 5 (cyan and magenta) and 6 (olive and blue) have the same luminance, and only one of these colors (cyan in Texture Set 5) is a target color. Thus, as we discussed in the introduction, the colors of the blobs in the synthesized textures differ in both luminance and chrominance, and hence differences in the contrast sensitivity of luminance and chrominance edges [24] cannot account for the results of our experiments.

The texture stimuli for our empirical studies contain nine sets of texture images that include posterized and synthetic images, cropped to 128×128 pixels and upsampled to 256×256 by pixel repetition. Note that Texture Sets 8 and 9 are the same as Texture Sets 2 and 7, respectively, but have different target colors.

3.1.4 *Texture Stimuli for Study 1: Posterized versus Varying Percentage Isotropic Blobs*

Each set includes one posterized texture and seven synthetic textures that share the same dominant colors with different

[1] http://www.corbis.com

percentages. Fig. 4(c) shows the color composition of the posterized textures of Fig. 4(b) with the target color underlined. The length of each color bar represents the percentage. Fig. 4(e) shows seven isotropic textures for each set, each of which was synthesized with target color percentage that differs by -9%, -6%, -3%, 0%, $+3\%$, $+6\%$, and $+9\%$ actual percentage points from the corresponding posterized texture. The percentages of the remaining colors were adjusted proportionately. Note that the posterized images of Texture Sets 8 and 9 are the same as those of Texture Sets 2 and 7, respectively, but the target colors are different, and so are the realizations of the synthetic textures. The target colors in Texture Sets 1–7 are the most illuminant colors, while the target colors in Texture Sets 8 and 9 are the most distinct colors.

With 8 textures (1 posterized and 7 synthetic), there are 28 possible image pairs in each set. Based on a preliminary test, when the actual difference in the amount of target color exceeds 9%, the difference is obvious and the participants are able to consistently select the image with the higher color amount. We thus only selected pairs with 9% or less color difference. This results in 15 synthetic–synthetic (*S–S*) pairs and 7 synthetic–posterized (*S–P*) pairs for each texture set. To balance the occurrence of S–S and S–P pairs, we repeated each S–P pair twice.

3.1.5 *Texture Stimuli for Study 2: Posterized versus Varying Scale Square Blocks*
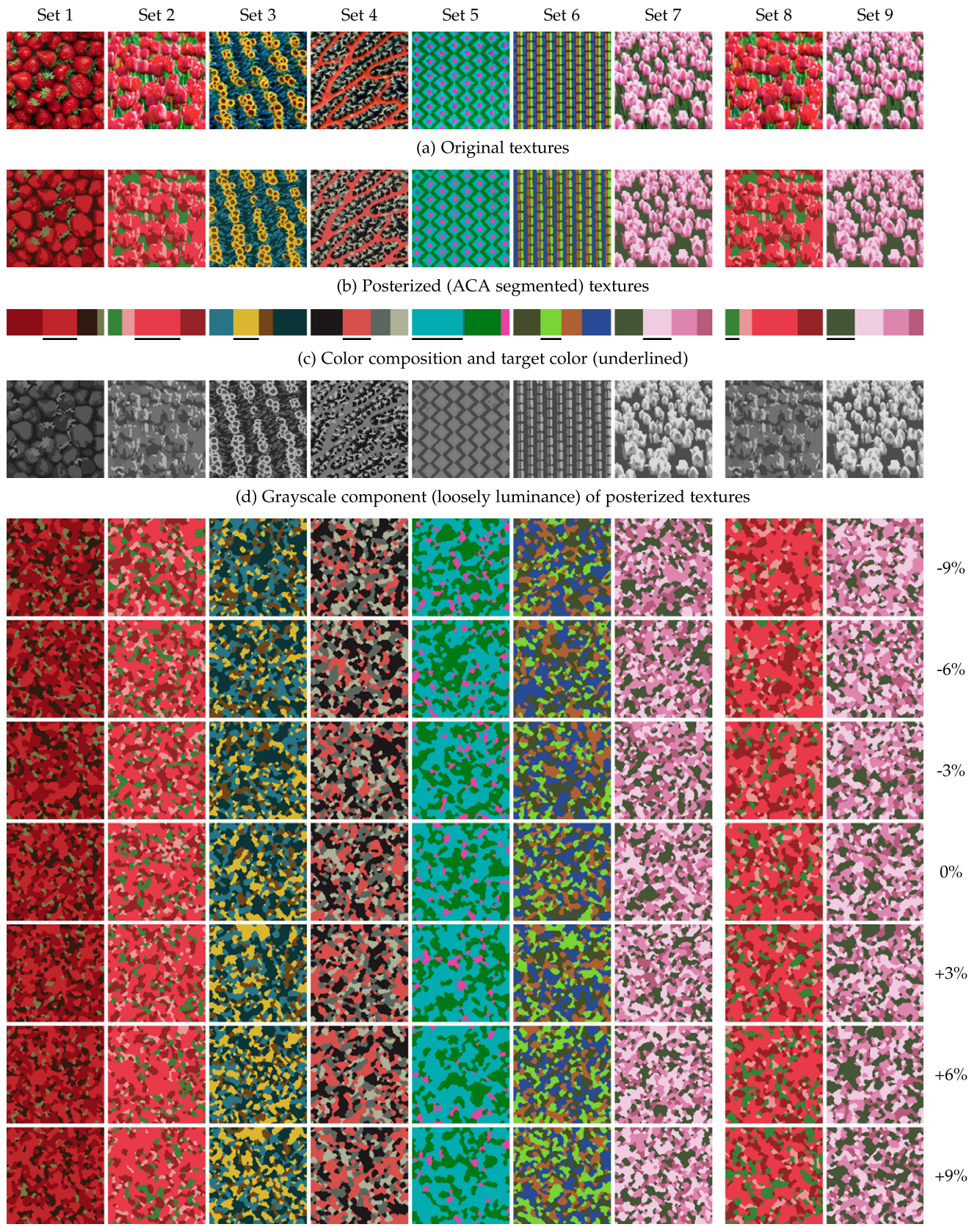
The texture stimuli contain nine sets of color texture images. Each set includes one posterized texture, the same as that used in Study 1, and four synthetic textures of varying scale that share the same color composition with the posterized texture. Fig. 5(a) shows the posterized textures, and Fig. 5(b) shows their color composition with the target color underlined. The synthetic textures were generated with squares of sizes 8×8, 16×16, 24×24, and 32×32, to obtain different scales, as shown in Fig. 5(c). The actual color composition of each synthetic image is the same as that of the corresponding posterized texture image.

However, to add more variability to the test patterns, we generated additional textures for each of the four scales of each color set with an additional 3% of the target color, while the other colors were adjusted proportionately. These textures are shown in Fig. 5(d).

Thus, in each set we have one posterized texture and a total of eight synthetic textures, two for each scale. In addition to gathering more data, we found that the added variability was necessary in order to stimulate the participants' interest, offering some relatively easy choices along with the more challenging ones, for which the participant might select one image at random. This keeps the participant alert and discourages reverting to random selection for all image pairs.

Note that in Fig. 5(c), even though the target colors are different, the synthetic images for Texture Sets 8 and 9 are the same as those of Texture Sets 2 and 7, respectively, because the color compositions are the same. However, in

(a) Original textures

(b) Posterized (ACA segmented) textures

(c) Color composition and target color (underlined)

(d) Grayscale component (loosely luminance) of posterized textures

(e) Isotropic synthetic images with varying target color amount (% target color amount difference from color composition in (c))

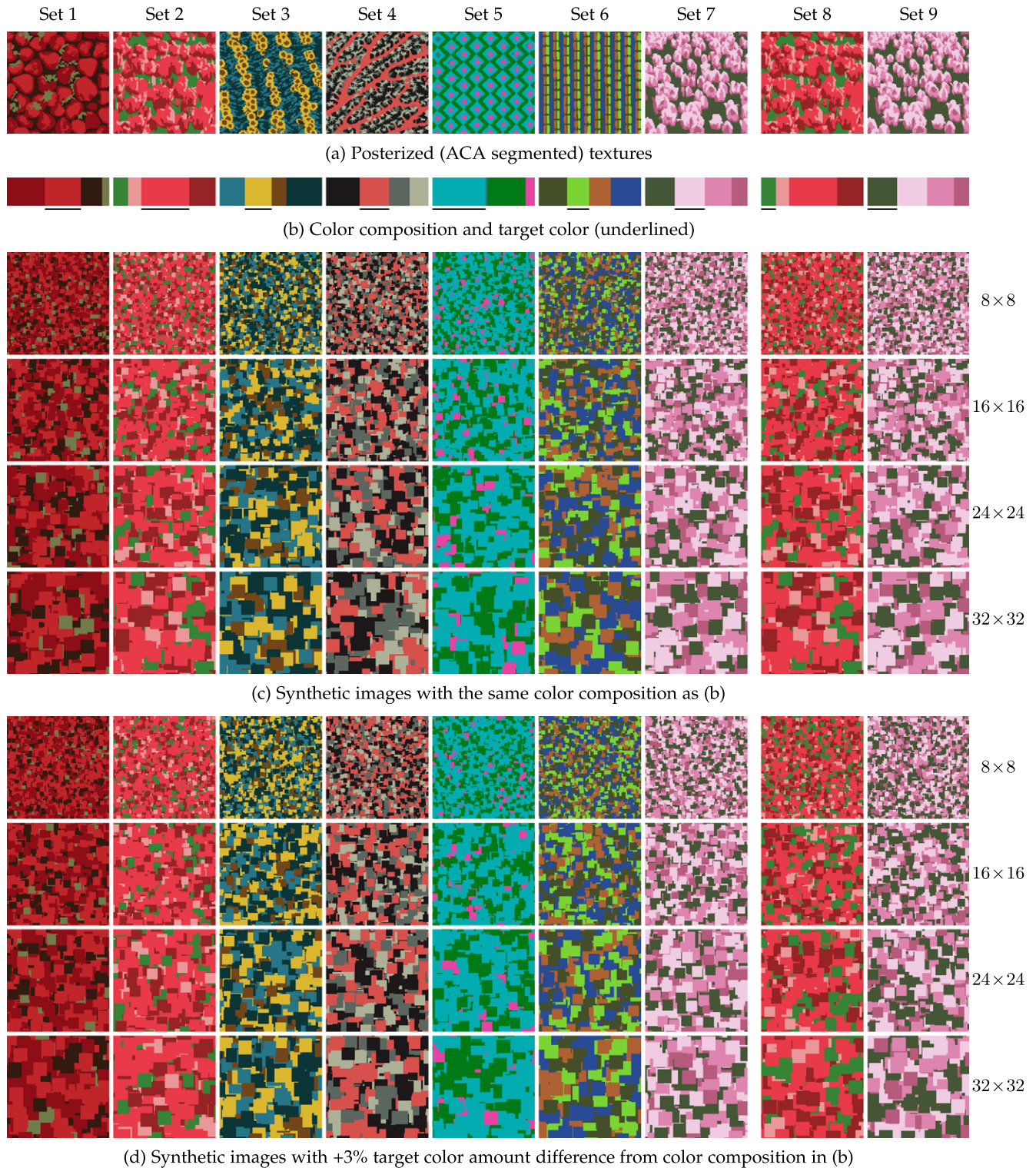Figure 4. Texture stimuli for Study 1.

(a) Posterized (ACA segmented) textures

(b) Color composition and target color (underlined)

(c) Synthetic images with the same color composition as (b)

(d) Synthetic images with +3% target color amount difference from color composition in (b)

Figure 5. Texture stimuli for Study 2.

Fig. 5(d), the color compositions are different and so are the realizations.

With 9 textures in each set (one posterized and 8 synthetic), there are 28 possible S–S and 8 S–P pairs for presentation to each participant. To balance the occurrence of S–S and S–P pairs, we repeated each S–P pair three times.

3.1.6 Texture Stimuli for Study 3: Varying Shape of Rectangular Blocks

The texture stimuli contain nine sets of color texture images. Each set includes four synthetic textures of varying shape that share the same dominant colors and percentages with the posterized texture of the corresponding set in Studies 1
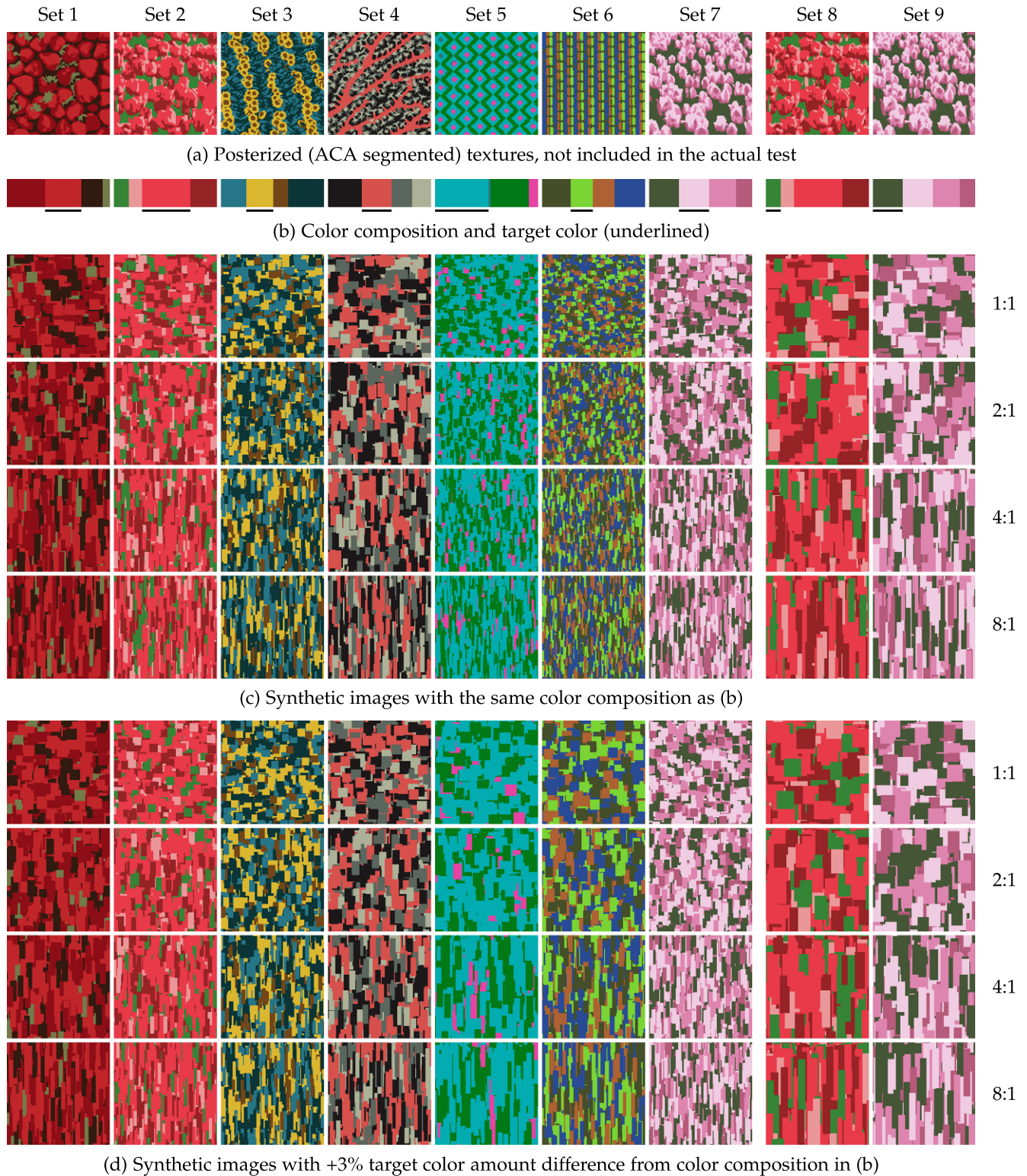
(a) Posterized (ACA segmented) textures, not included in the actual test

(b) Color composition and target color (underlined)

(c) Synthetic images with the same color composition as (b)

(d) Synthetic images with +3% target color amount difference from color composition in (b)

Figure 6. Examples of texture stimuli for Study 3.

and 2. The posterized textures were not included in this study. The synthetic textures were generated with rectangles with H/W ratios 1 : 1, 2 : 1, 4 : 1, and 8 : 1, that is, with square and rectangular blocks with different elongation factors. As we discussed, all of the rectangular blocks have the same area in order to maintain a constant scale. Note that, in contrast

to the other two studies, the size of the rectangular block for each set was based on the estimated scale of the target color in the posterized texture; the scale estimation algorithm is presented in Appexdix B. Figure 6(a) shows the posterized textures, and Fig. 6(b) shows their color composition with the target color underlined. Fig. 6(c) shows the synthetic textures
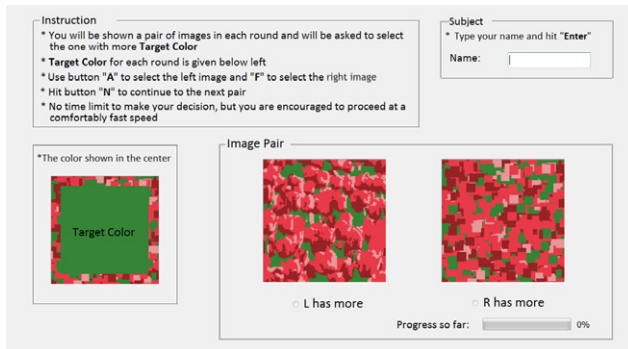
Figure 7. Graphical user interface.

with the color composition shown in Fig. 6(b). As in the varying scale study, for each H/W ratio, we also generated textures with an additional 3% of the target color, shown in Fig. 6(d), for more variability. Note that, since the scale of each synthetic image is based on the estimated scale of the target color in the posterized texture, the scales of Texture Sets 2 and 8 are different; the same is true for Texture Sets 7 and 9.

With 8 textures in each set (all synthetic), there are 28 possible pairs for presentation to each participant.

3.1.7 *Procedure*
The empirical studies were designed as 2AFC tests. The graphical user interface is shown in Figure 7. A sequence of image pairs was presented to the participants side by side. Each image pair contained textures that had the same dominant colors (corresponding to the same posterized texture) but differed in texture structure and/or target color percentage. The positions of the images in a pair were randomized in each trial to eliminate response bias. The participants were asked to select the image that contained a higher percentage of the target color using keyboard shortcuts.

The three studies (isotropic blobs, square textures of different scales, and textures with different rectangle shapes) were conducted separately. To prevent any participant biases based on parameter values (e.g., blob size or shape), image pairs that correspond to the nine different posterized images were mixed up and presented in random order. There were no time limits in any of the tests. However, the participants were encouraged to proceed at a comfortably fast pace.

4. ANALYSIS OF THE RESULTS
4.1 *Analysis of Study 1*
The first study was carried out with the posterized textures and the synthetic isotropic blob textures. As we discussed, the goal of this study was to determine whether the difference in structure between the posterized texture and the synthetic isotropic blob textures affects the perception of the target color amount. As we will see, this study also demonstrates that the perceived differences in the amount of target color are consistent with the actual color percentages.

We employed Thurstonian scaling [54] to convert the 2AFC results to preference scores for each texture set. The model assumes that the relative magnitudes of the preferences for the stimuli can be determined by the winning frequencies that one stimulus is selected over another in a paired comparison task. We accumulated the comparative choices across participants for each pair and computed a preference matrix for each texture set with the winning frequencies between 0 and 1. The values in each preference matrix were omitted and treated as missing values when the winning frequency was too small (< 0.02) or too large (> 0.98) to give stable estimates [11]. The diagonal values were set to 0.5. We then applied the Thurstone Case V model to convert pairwise preferences to continuous perception scores. The scores were further normalized to Z-scores to facilitate evaluation across the texture sets. The Z-scores are obtained by subtracting the mean and dividing by the standard deviation. Larger Z-score indicates stronger perception of the target color amount.

Figure 8 displays the Z-scores of each texture set as a function of the physical target color percentage difference. The dotted line in each plot represents the perception of the posterized image among the synthetic images. The overall increasing tendencies in Fig. 8 show that the participants are capable of perceiving the changes in the amount of target color. The fitted linear regression between the Z-scores and the physical target color percentage difference in each texture set suggests that the perceived color amount is linearly related to the physical percentage difference in target color. We counted the correct responses of the 2AFC results between S–S pairs. The probabilities of correct response for 3%, 6%, and 9% target color differences are 0.82, 0.9, and 0.98, with 36, 45, and 36 samples, respectively, which indicates that the just noticeable difference of color amount is below the 3% difference in color.

By comparing the Z-scores of the posterized and the synthetic images, we notice that the perceived target color amount of most texture sets is between the perceived target color amount of the synthetic images with $\pm3\%$ actual target color amount difference. The only exceptions are Texture Sets 1 and 5. For Texture Set 1, the posterized texture is perceived as having a much higher percentage of the target color than the synthetic image with 0% target color difference, while the opposite is true for Texture Set 5. Table I lists the perceived target color difference (ΔZ) of each texture set.

We now examine the factors that may have caused the perceived color amount difference between the posterized and the synthetic image (0%). Three possible factors are explored: (a) the scale difference of the target color blobs in the posterized and synthetic images (ΔS); (b) the lightness difference between the posterized and the synthetic images (ΔL); (c) the color distance between the target color and the most similar adjacent color (ΔE).

One obvious difference is the color scale. For Texture Set 1 of Fig. 4, the target color blobs of the posterized texture appear to be larger than those of the synthetic textures, while the opposite is true for Texture Set 5. To quantitatively check
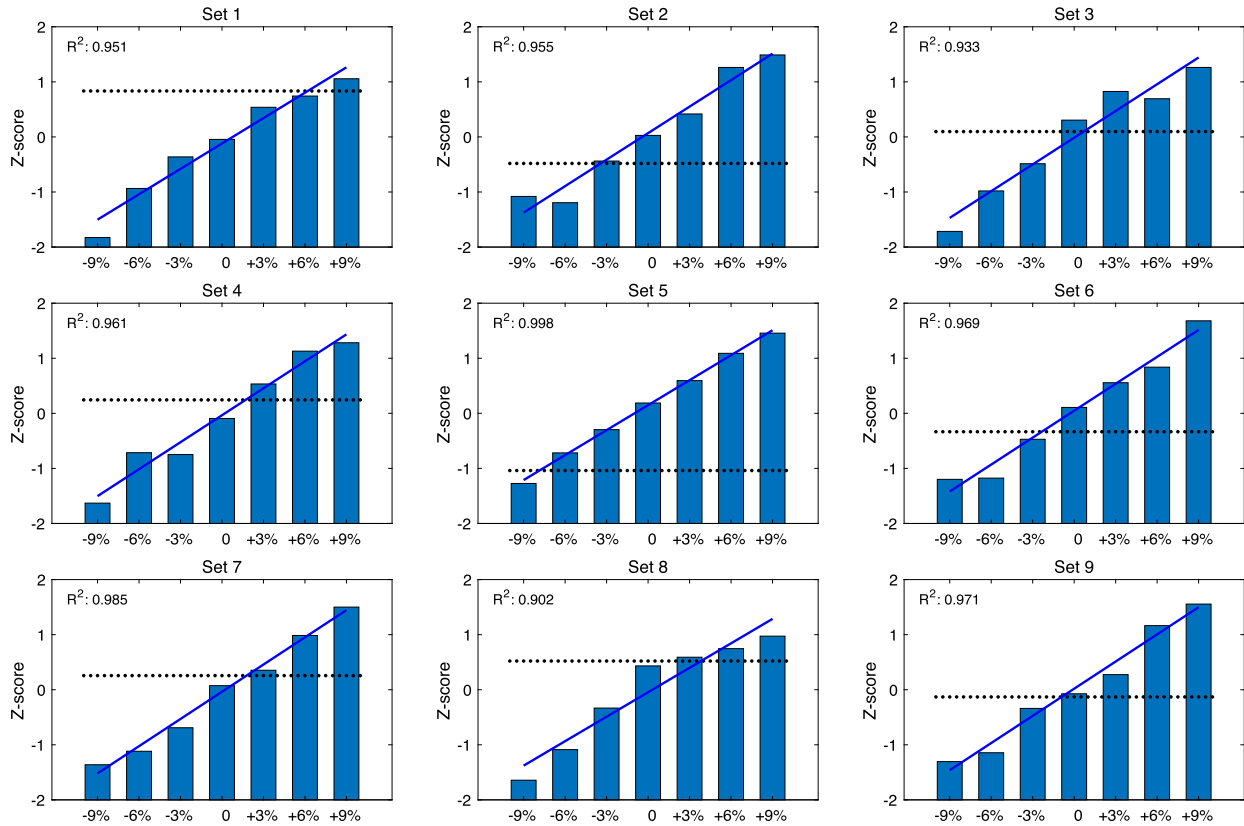
Figure 8. Analysis of the results of Study 1. Ranking the preference scores to generate the perception of a larger amount of color for synthetic textures with varying actual target color amount. The x-axis represents the target color amount difference (synthetic–posterized). The y-axis represents the Z-score for each image using Thurstonian scaling. The dotted lines correspond to the posterized texture images in each texture set. The solid blue lines are linear regressions fitted to the synthetic images of each texture set. The R^2 values of the fitted regressions are shown in the upper left corner.

Table I. Perceived color amount difference between the posterized and the synthetic (0%) textures and the corresponding difference in scale, color, and lightness.

Posterized–Synthetic (0%)		Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9
Z-score	(ΔZ)	0.879	−0.509	−0.208	0.337	−1.226	−0.441	0.183	0.089	−0.057
Scale	(ΔS)	4.98	−2.83	−2.36	−1.41	−13.3	−4.65	−2.01	2.78	2.01
Adjacent color	(ΔE)	17.0	27.4	55.6	64.2	60.5	75.9	30.8	78.5	58.2
Lightness	(ΔL)	0.061	0.159	0.905	0.737	0.039	−0.250	0.186	0.151	0.313
Fitted regression		$\Delta Z = c_1 \cdot \Delta S + c_2 \cdot \Delta L + c_3 \cdot \Delta E + \text{interp}$								
	Coefficients	$c_1 = 0.093$			$c_2 = 0.16$			$c_3 = -0.004$		interp $= 0.26$
		p-value: 0.009			p-value: 0.64			p-value: 0.45		p-value: 0.43
	Standardized	0.83			0.10			0.16		
	Fitted regression has $R^2 = 0.81$, p-value $= 0.029$									

the above observations, we calculated the *average linear scale* of the target color blobs in each image as the geometric mean of the average horizontal and vertical lengths of the blobs, as described in appendix B. Table I lists the scale difference (ΔS) between the texture synthesized with 0% difference in the target color and the posterized texture. To check how the scale difference relates to the perceived target color amount difference, we fitted a linear regression between ΔZ and ΔS ($\Delta Z = c \cdot \Delta S + \text{interp}$). The estimated coefficients ($c = 0.1$, $p = 0.002$, $\text{interp} = 0.08$) with $R^2 = 0.778$ show that the scale

difference has high positive correlation with the perceptual difference.

Another factor that may have influenced the outcome of our experiments is the average lightness of the image. Note that if the target color is lighter or darker than the other colors, then by modifying the amount of target color we are also modifying the average lightness of the image. Thus, one could hypothesize that the participants may have ordered the textures according to overall lightness rather than according to the amount of perceived target color,
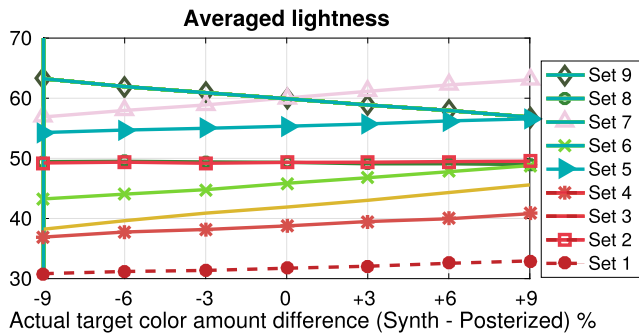
Figure 9. Average lightness of different synthetic images.

as instructed. Figure 9 plots the average lightness of each synthetic image for each texture set. Note that the most dramatic changes occur in Texture Sets 3, 4, 6, 7 (increasing slope), and 9 (decreasing slope). However, the placement of the posterized texture does not seem to be related to the change in average lightness. The lightness difference ΔL between the posterized and the synthetic (0%) images for each texture set is listed in Table I. Similar to the analysis of scale difference, we fitted a linear regression between ΔZ and ΔL ($\Delta Z = c \cdot \Delta L + \text{interp}$) to check the relation between the perception and lightness. The estimated coefficients ($c = 0.36$, $p = 0.57$, interp $= -0.198$) with $R^2 = 0.047$ show that the lightness does not play a significant role in influencing the perceived difference in color amount.

Apart from the color scale and lightness, the similarity of spatially adjacent colors could also affect the perceived target color amount, especially for Texture Set 1. In Texture Set 1, the most similar colors (two shades of strawberries) are adjacent in the posterized texture but randomly placed in the synthetic texture. ΔE in Table I lists the distance between the target color and the most similar adjacent color for each texture set. Note that the most similar adjacent color (dark red) occurs for Texture Set 1; for all the other sets, the difference between adjacent colors is quite large, and hence does not seem to have any effect on the perceived target color amount.

To investigate the relative contributions of each factor, we fitted a multiple regression with the three factors on the perceived color amount difference ΔZ ($\Delta Z = c_1 \cdot \Delta S + c_2 \cdot \Delta L + c_3 \cdot \Delta E + \text{interp}$). The estimated coefficients with the corresponding p-values and standardized coefficients are listed in Table I. Note that ΔS and ΔL contribute positively to ΔZ, and ΔE negatively, but the p-values and the standardized coefficients show that only the ΔS contributions to ΔZ are statistically significant, and that the contributions of the other independent variables are negligible.

In the analysis of Study 2 below, where we vary the scale of the blobs for a fixed color composition (and hence fixed lightness), we further investigate the relationship between the target color scale and the perception of target color amount.

4.2 *Analysis of Study 2*

The goal of this study was to consider the direct effects of scale on the perceived amount of the target color. This study was carried out with the posterized textures and textures synthesized with squares of different sizes. We compared the synthetic textures to each other, as well as to the posterized texture. As shown in Fig. 5, each color set has four different scales (generating block sizes of 8×8, 16×16, 24×24, and 32×32) with the same color composition, and the same four scales with +3% of the target color.

As in Study 1, we employed Thurstonian scaling [54] to convert the 2AFC results to preference scores for each texture set. Using the nine sets as replication, a repeated-measures analysis of variance (ANOVA) with two within-subject factors (block scale level and percentage) verified that the perceived target color amount is affected by the scale (block size), $F(3, 24) = 123.4$, $p < 0.0001$, and by the actual color amount change, $F(1, 24) = 109.3$, $p < 0.0001$. There is no interaction between the block size and actual color amount, $F(3, 24) = 0.578$, $p = 0.635$.

Figure 10 shows the Z-scores of each texture set. A higher score indicates a higher amount of perceived target color. It is no surprise that the values for the +3% images are higher than those for the 0% images at each scale since the perceived color amount is consistent with the actual color amount, as shown in Study 1. In addition, the increasing scores for each of the texture sets in Fig. 10 demonstrate that pattern agglomeration (increasing scale) results in the perception of increasing target color amount. The score of the posterized texture is shown in dotted black line for each texture set. It is clear that the placement of the posterized images varies considerably across the texture sets.

Figure 11 presents the average linear scale of the target color blobs of the posterized and synthetic textures. A two-way ANOVA was employed to test whether, in addition to the generating block size, the +3% color difference affected the scale values. The analysis shows that apart from the statistically significant effect of block size ($F(3, 24) = 73.9$, $p < 0.0001$), the average linear scale is also affected by the actual color percentage difference ($F(1, 24) = 5.84$, $p = 0.02$). There is no interaction between the block size and actual color amount, $F(3, 24) = 0.29$, $p = 0.83$.

Note that for Texture Sets 1, 4, 8, and 9, the linear scale of the posterized texture is relatively higher compared to that of the synthetic textures, and the scale of Texture Set 5 relatively lower, which is consistent with the results in Fig. 10. Note, however, that the relative location of the posterized (dotted black line) compared to the synthetic textures in Fig. 10 is generally higher than the corresponding location of the posterized relative to the synthetic textures in Fig. 11. Here we should also point out that even though the images synthesized with a given generating square size are expected to have the same scale, and hence the 0% and +3% bar heights are expected to coincide, the actual estimated scales differ due to the varying percentages (as we noted
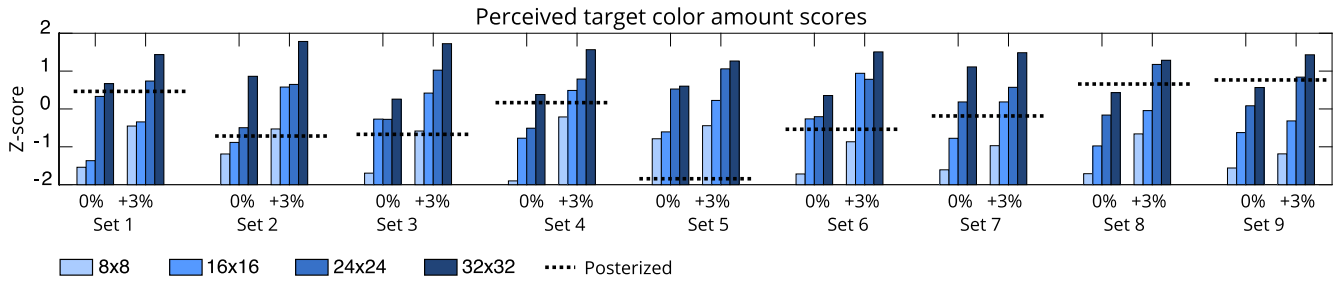
Figure 10. Analysis of the results of Study 2. Ranking the preference scores to generate the perception of a larger amount of color for synthetic textures with varying pattern scales—with the same color composition as the posterized texture (0%) and +3% target color—and posterized texture (dotted line).
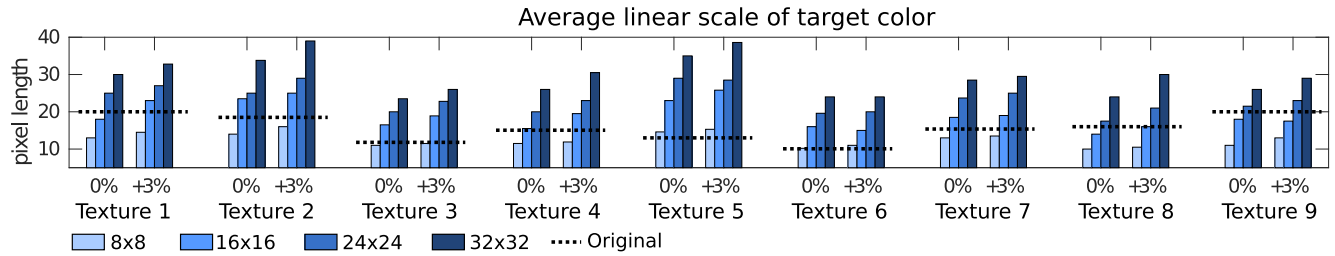

Figure 11. Analysis of the results of Study 2. Estimated average scale of target color in each image—with the same color composition as the posterized texture (0%) and +3% target color–and posterized (dotted line).

above, higher percentages result in higher probability of agglomeration) and the randomness in the block placements.

To summarize the relationship between color scale and color amount perception, Figure 12 shows a combined view of Figs. 10 and 11 across all texture sets (posterized textures excluded). The position of each dot is obtained by averaging the linear scales (x-axis) and the Z-scores (y-axis) of all synthetic textures with the same color block scale (8×8, 16×16, 24×24, 32×32) and the same actual color amount difference (0%, +3%). We used linear regression ($Z = b \cdot$ Scale $+ \delta \cdot$ ColorAmountDifference $+$ interp) to investigate the influence of the average linear scale on the perceived color amount (Z-score) with the actual color amount difference as a dummy variable. The estimated coefficients (slope: $b = 0.102$, $p < 0.001$; $\delta = 0.709$, $p < 0.001$; interp $= -2.48$, $p < 0.001$) demonstrate that there exists a linear positive relationship between the target linear scale and the perceived color amount.

4.3 *Analysis of Study 3*

The goal of this study was to investigate the effect of shape on the perception of the amount of the target color. In this study, we only compared synthesized textures (with rectangles of different H/W ratios) to each other. The analysis is similar to that of the second study. Using the nine textures as replication, a repeated-measures ANOVA with two within-subject factors (elongation level and percentage) verified that the perceived target color amount is affected by the H/W ratio, $F(3, 24) = 16.46$, $p < 0.0001$, and by the actual color amount change, $F(1, 24) = 190.8$, $p < 0.0001$. The actual color amount and elongation degree in structure are independent with no interaction, $F(3, 24) = 0.654$, $p = 0.589$.
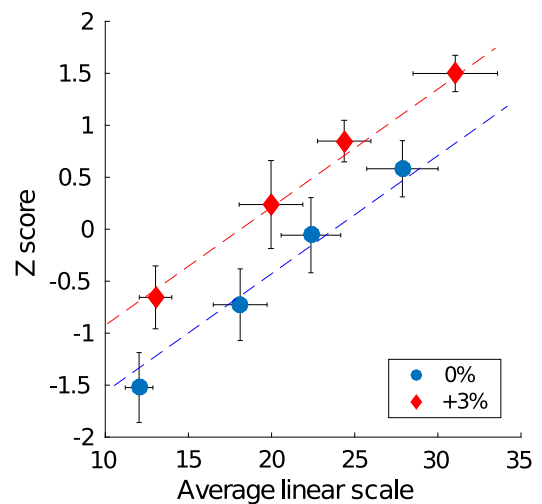

Figure 12. Effects of average linear scale on the Z-scores of the perceived color amount. Error bars are the standard error of the mean score and average linear scale across all texture sets. The dashed lines are linear regression lines fitted to all the data points with the actual color amount as a dummy variable.

Figure 13 shows the estimates of the perceived scores within each color set. The labels of the x-axis are the H/W ratios of color blocks used to synthesize each texture. The +3% synthesized textures are almost always higher than the 0% synthesized textures at each H/W ratio due to the increase in the actual target color amount. Apart from Texture Set 6, the general trend is that the perceived amount of target color decreases as the geometric shape becomes more elongated (H/W ratio increases) for both the 0% and the +3% textures. One possible explanation of the weak relation between shape
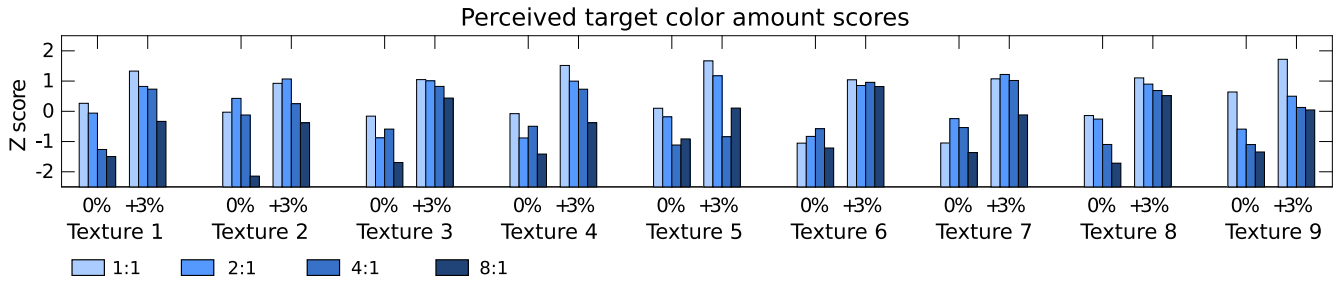
Figure 13. Analysis of the results of Study 3. Ranking the preference scores to generate the perception of a larger amount of color for synthetic textures with varying pattern shapes—with the same color composition as the posterized texture (0%) and +3% target color.
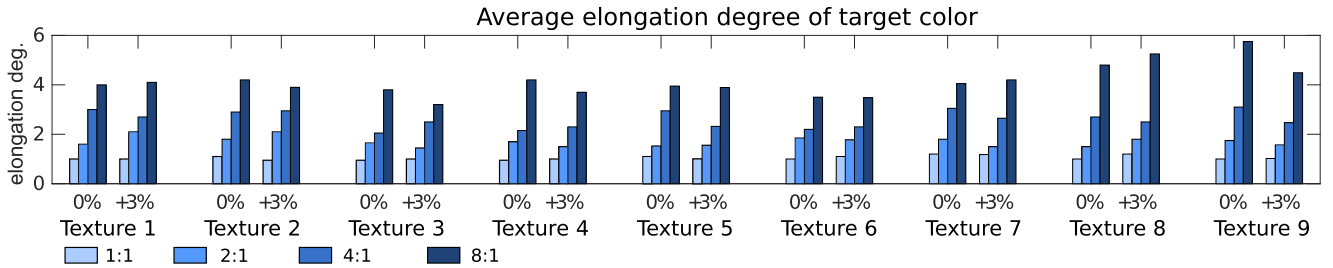

Figure 14. Analysis of the results of Study 3. Estimated average elongation of target color in each image—with the same color composition as the posterized texture (0%) and +3% target color.

and color perception in Texture Set 6 is the influence of pattern scale, which is considerably smaller than in the other sets. While the H/W ratio was changed, the scales of all synthesized textures in this study were designed with scales similar to the posterized texture. When the pattern scale is small, the shape variations are not as noticeable as they are in the larger pattern scales.

Figure 14 shows the bar plots of the average elongation degree of each texture composed of color blocks with the specified H/W ratio. The average elongation degree is defined as $\frac{\max(L_v, L_h)}{\min(L_v, L_h)}$, where L_v and L_h are the average vertical and horizontal lengths of the target color blobs based on the algorithm in Appendix A. As expected, the elongation degree increases as the color block H/W ratio increases. Note, however, that the resulting elongation degree is not equal to the H/W ratio, due to block overlap and random block placement. As in the analysis of the scale study, a two-way ANOVA was employed to test whether, in addition to the generating block H/W ratio, the +3% color difference affected the elongation degree values. The analysis shows that apart from the statistically significant effect of color block H/W ratio ($F(3, 24) = 231, p < 0.0001$), the influence of actual color percentage difference on the average elongation degree is not statistically significant ($F = 0.97, p = 0.33$). There is no interaction between the color block H/W ratio and the actual color amount, $F(3, 24) = 0.5, p = 0.68$.

To summarize the relationship between color elongation and color amount perception, Figure 15 shows a combined view of Figs. 13 and 14 across all texture sets. The position of each dot represents the average elongation degree (x-axis) and the Z-score (y-axis) of

all synthetic textures with the same color block ratio (1 : 1, 2 : 1, 4 : 1, 8 : 1) and the same actual color amount difference (0%, +3%). We used linear regression ($Z = b \cdot \text{Elongation} + \delta \cdot \text{ColorAmountDifference} + \text{interp}$) to investigate the influence of shape (average elongation degree) on the perceived color amount (Z-score) with the actual color amount difference as a dummy variable. The estimated coefficients (slope: $b = -0.38, p < 0.001$; $\delta = 1.366, p < 0.001$; $\text{interp} = 0.217, p = 0.08$) demonstrate that there exists a linear negative relationship between the elongation degree and the perceived color amount.

5. MODELING THE RESULTS
An important question to be addressed is whether there is a model that can explain and predict the results of our empirical studies. While it would be desirable to come up with a model of the underlying visual mechanisms that explain the observed effects, in this section we propose a model that estimates the contribution of each pixel to the perceived amount of the target color, as determined by our empirical studies. We hope that this will stimulate interest in exploring the underlying visual mechanisms.

Our hypothesis is that there is an edge effect, whereby the pixels near the blob boundaries contribute less to the perceived target color amount than pixels at the center of the blobs. That is, the visibility of a pixel decreases as the pixel approaches the closest boundary. We will assume that this effect is limited to a distance of a few pixels from the boundary.
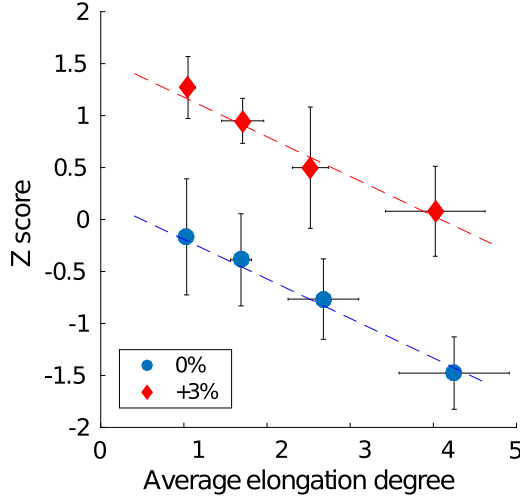
Figure 15. Effects of average elongation degree on the Z-scores of perceived color amount. Error bars are the standard error of the mean score and average elongation degree across all texture sets. The dashed lines are linear regression lines fitted to all the data points with the actual color amount as a dummy variable.

To estimate the distance of each pixel from the closest boundary, we rely on mathematical morphology [33]. To obtain the pixels that are adjacent to a color boundary, we apply a one-pixel erosion operator to the map of the target color. The eroded pixels comprise the first layer of pixels that are at a distance of one pixel from a boundary. We then apply a one-pixel erosion to the remaining pixels to obtain the second layer of pixels that are at a distance of two pixels from a boundary. We can continue this process to obtain multiple layers at different distances from the boundaries. Figure 17 shows an example of the iterative erosion method; Fig. 17(a) shows a color segment from one of the synthesized images, Fig. 17(b) shows the decomposition into layers, and Fig. 17(c) shows the number of pixels in each layer.

Let $P(t)$ represent the perceived proportion of a texture image that is taken by a given target color t. Assume there are K boundary layers in total, the first being the outermost layer of each color segment and the Kth being the innermost layer of the largest segment. Then $P(t)$ is given by the layer-weighted area of the target color

$$P(t) = \frac{1}{N_t} \sum_{k=1}^{K} w_k N_{t,k}, \qquad (2)$$

where $N_{t,k}$ is the number of pixels in a layer that is at a distance of k pixels from the closest boundary, N_t is the total number of pixels of the target color t, and w_k is a weight associated with the visibility of the kth layer.

For the layer weights, we tried a hyperbolic tangent function

$$w_k = \tanh(a \cdot k + b) = \frac{e^{(a \cdot k + b)} - e^{-(a \cdot k + b)}}{e^{(a \cdot k + b)} + e^{-(a \cdot k + b)}}, \ a+b \geq 0, \ a > 0, \qquad (3)$$

where k is the layer index, w_k takes values in the interval $[0, 1]$, and a and b are constants controlling the saturation

speed and shift of w_k, respectively. With large a, the weights saturate quickly to 1. Using $a = 0.5$, $b = -0.25$, we obtained weights 0.24, 0.63, and 0.85 for the outermost layers, with the weights for the remaining inner layers taking values close to 1. In a limiting case, $w_1 = 0$ and $w_k = 1$ for $k > 1$ ($a = 5$, $b = -5$). The reason for these choices is that visibility is reduced near the boundaries and increases rapidly as we move away from the boundaries. Of course, all this depends on the viewing distance and display resolution. As we discussed in the introduction, the viewing distance and display resolution in our experiments are such that there is very little filtering across the boundaries. Finally, we also tried linearly increasing weights as well as the default constant weights $w_k = 1$ for all k (no edge effect). Three different weighting strategies are illustrated in Figure 16.

Tables II and III show the Pearson correlation coefficient between the model estimates and the perceived target color amount (Z-scores), for each texture set, for the scale and shape studies, respectively. We also calculated the average correlation coefficient across all texture sets as the inverse transform of the average of Fisher's Z-transform of each texture set [51]. The tables compare three parameter settings of the hyperbolic weighting strategy with linearly increasing weights and the default constant weight strategy ($w_k = 1$ for all k).

Table II shows that, for the scale study, all the weighting strategies provide acceptable performance, with the linear weights outperforming the other strategies, and demonstrates the inadequacy of the constant weights. On the other hand, for the shape study, Table III shows that even though the linear weights have higher overall correlation than the hyperbolic weighting schemes, they have relatively low correlation on Texture Set 6 (0.666). Overall, the quick saturation weighting strategy using a hyperbolic weighting scheme with $a = 0.5$ and $b = -0.25$ provides a relatively stable strong positive correlation with the perceived color amount, for the best overall performance. However, the performance of the limiting case with $w_1 = 0$ and $w_k = 1$ for $k > 1$ is almost as good.

Using the hyperbolic weighting scheme with $a = 0.5$, $b = -0.25$, we can estimate the perceived color amount for each image in the two studies. The results are plotted in Figure 18 as a function of the measured Z-scores. The plots provide another indication that the model is consistent with the results of our empirical studies.

As we discussed above, we could not find any visual mechanisms that explain the observed edge effect. The closest is a study by Eskew and Boynton [13], who consider compact, spatially localized stimuli with varying width and junction length; however, the conclusions are about variations in contrast sensitivity, which is not likely to have an effect on stimuli with contrast well above threshold.

6. DISCUSSION AND CONCLUSIONS
We investigated the influence of texture structure on the perception of color composition through a series of empirical studies. We relied on color segmentation (adaptive

Table II. Scale study: Pearson correlation between layer-weighted area of target color and perceived target color amount.

Weight Strategy	Pearson	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Average
$a = 0.5, b = -0.25$	Correlation	0.896	0.911	0.982	0.938	0.919	0.958	0.950	0.932	0.888	0.938
	p-value	0.0026	0.0017	0.0001	0.0006	0.0012	0.0002	0.0003	0.0008	0.0032	–
$a = 0.5, b = 0$	Correlation	0.838	0.881	0.934	0.938	0.931	0.954	0.938	0.850	0.779	0.907
	p-value	0.0093	0.0039	0.0007	0.0006	0.0008	0.0002	0.0006	0.0075	0.0229	–
$w_1 = 0, w_k = 1$ for $k > 1$	Correlation	0.882	0.893	0.978	0.933	0.918	0.957	0.942	0.935	0.876	0.932
	p-value	0.0037	0.0029	0.0000	0.0007	0.0013	0.0002	0.0005	0.0007	0.0044	–
Linear weights	Correlation	0.944	0.945	0.977	0.902	0.923	0.915	0.963	0.968	0.983	0.954
	p-value	0.0004	0.0004	0.0000	0.0022	0.0011	0.0015	0.0001	0.0001	0.0000	–
$w_k = 1$ for all k	Correlation	0.401	0.541	0.580	0.633	0.448	0.501	0.404	0.516	0.328	0.489
	p-value	0.3243	0.1663	0.1325	0.0918	0.2661	0.2060	0.3215	0.1905	0.4284	–

Table III. Shape study: Pearson correlation between layer-weighted area of target color and perceived target color amount.

Weight Strategy	Pearson	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Average
$a = 0.5, b = -0.25$	Correlation	0.852	0.750	0.856	0.887	0.881	0.749	0.942	0.750	0.720	0.837
	p-value	0.0072	0.0321	0.0066	0.0033	0.0038	0.0361	0.0005	0.0321	0.0440	–
$a = 0.5, b = 0$	Correlation	0.812	0.679	0.824	0.860	0.865	0.845	0.934	0.742	0.713	0.823
	p-value	0.0143	0.0641	0.0120	0.0062	0.0055	0.0083	0.0007	0.0349	0.0473	–
$w_1 = 0, w_k = 1$ for $k > 1$	Correlation	0.841	0.730	0.861	0.882	0.866	0.720	0.942	0.742	0.725	0.830
	p-value	0.0089	0.0398	0.0060	0.0038	0.0055	0.0438	0.0005	0.0350	0.0417	–
Linear weights	Correlation	0.940	0.864	0.892	0.917	0.906	**0.666**	0.894	0.815	0.749	0.868
	p-value	0.0005	0.0056	0.0029	0.0014	0.0019	0.0717	0.0027	0.0136	0.0325	–
$w_k = 1$ for all k	Correlation	0.700	0.486	0.742	0.784	0.821	0.958	0.875	0.713	0.670	0.786
	p-value	0.0530	0.2219	0.0350	0.0213	0.0125	0.0002	0.0044	0.0472	0.0693	–
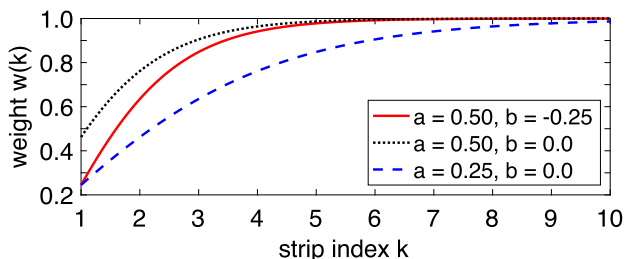
Figure 16. Weights for kth layer using different strategies.

clustering) to estimate the color composition of a given texture image, and synthesized new textures with the same color composition but different structures, consisting of isotropic blobs and geometric blocks of different scales and shapes.

The first observation is that when scale and shape are not changing significantly, there is a linear relation between the actual color amount and the values derived from the 2AFC experiments. This suggests that the participants are able to consistently assess differences in color composition for textures of similar shape and scale. Second, the results of our empirical studies indicate that pattern scale and shape have a

strong impact on the perception of the target color amount. In particular, when images have the same physical amount of a target color, there exists a positive linear relationship between the Z-scores of the perceived color amount and the average linear scale of the target color blobs, and a negative linear relationship between the Z-scores of the perceived color amount and the average elongation degree of the target color blobs. However, we found one indication that the elongation effect might be weakened as the texture scale decreases (Texture Set 6 in Study 3). In addition, we found that there is no interaction between the actual color amount and the scale or the shape.

As we discussed in the introduction and the design of the empirical studies, we did not try to separate the effects of luminance and chrominance. With only a couple of exceptions, the colors of the blobs differ in both luminance and chrominance so that differences in contrast sensitivity of luminance and chrominance edges cannot account for the observed results. Selecting isochrominant stimuli would have led to very similar results, but would have made the stimuli less realistic and less appealing to the participants. In addition to visual appeal, selecting isoluminant stimuli could have induced averaging across the blob boundaries, which we
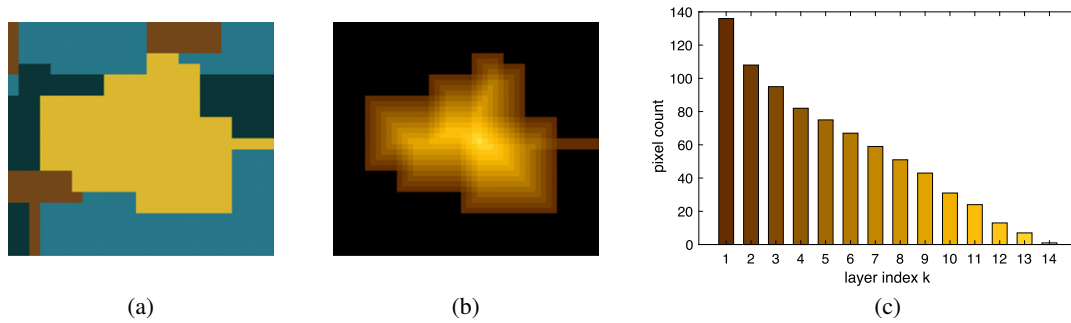
Figure 17. Iterative 1-pixel erosion of a target color segment. (a) Synthetic image. (b) Decomposition of target color segment into layers. (c) Distribution of layer pixel counts.
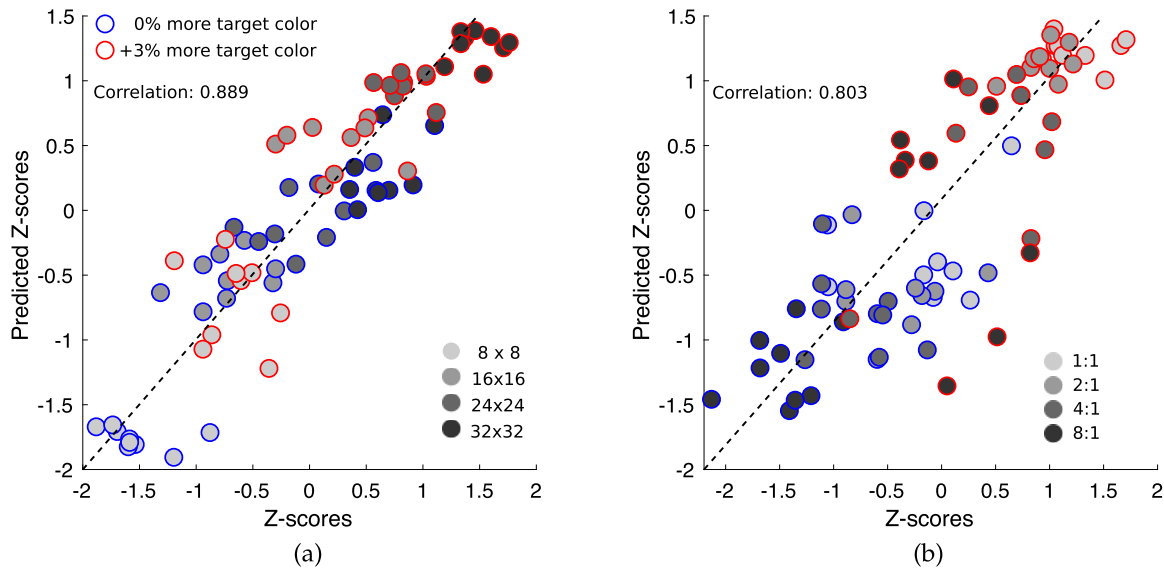

Figure 18. Model estimates versus measured Z-scores of perceived color amount for textures in (a) scale study and (b) shape study. The root mean squared error of the estimates for the scale study is 0.38 and for the shape study is 0.45. The dashed lines are fitted regression lines with slope 1 and intercept 0.

wanted to avoid. Our analysis showed that lightness did not play a significant role in influencing the perceived difference in color amount, which post hoc justifies our selections.

Based on our conclusions, the perceived amount of target color depends on the physical color amount, the scale, and the elongation degree. In Figs. 12 and 15, we showed how the perceived target color amount depends on estimates of scale and elongation, respectively. However, we also found that a relatively simple model that accounts for decreased visibility of the pixels near color boundaries is consistent with the results of our empirical studies.

A somewhat surprising finding is that the increase in the perceived color amount with increasing scale and decreasing elongation is true for two different target colors, the most illuminant and the most distinct color. If the same were true for the other two colors (four in total), the apparent paradox is that the percentages would add up to more than 100%. Alternatively, if the perceived color amount of each color decreases with decreasing scale and increasing elongation, and this were true for all colors, then the percentages

would add up to less than 100%, that is, the participants underestimate the amount of each color. We believe that the latter is the case and that the participants treat the rest of the colors as clutter, and apparently, the pixels with decreased visibility near the borders are included in the clutter.

The above finding is a consequence of the fact that we asked the participants to focus on one color. Our empirical studies are essentially equivalent to asking the participants to estimate the percentage of one color, effectively treating all other colors as clutter. If instead we had asked the participants to simultaneously estimate the percentage of each color, we expect that they would have produced numbers that add up to 100%. The testing of this hypothesis is beyond the scope of this article.

As we discussed in the introduction, the results of our empirical studies are consistent with Julesz's texton theory [20], whereby the size and shape of the textons affect the texture perception. We also found some interesting analogies with well-studied (but not directly related) phenomena, such as texture metamers [8]. Apart from the pixel visibility model

we proposed in this article, we could not find any other mechanisms that can explain the results of our studies, which have been motivated from engineering applications and have relied on real-world images that are more complex than typical stimuli used in psychophysical studies. However, we hope that our work raises new challenges for vision research, including the need for a more basic understanding of dominant colors of textures and general images.

Our results have definite implications for texture similarity metrics. If we assume that scale and elongation affect all the dominant colors, any metric adjustments should depend on estimates of the relative scale and elongation degree of each color of the images being compared. Better yet, the metric adjustments should depend on the model we proposed. As we discussed, according to the results of our empirical studies and the model estimates, the perceived color amounts may add up to less than the total texture area. As the scale of the images we compare decreases and the elongation increases, the perceived amount of the target color, and presumably the other colors, decreases. One could then argue that color composition differences should be given less weight than structure differences in the overall evaluation of texture similarity. The specifics of such adjustments will be the topic of future research.

APPENDIX A. TEXTURE SYNTHESIS ALGORITHMS

A.1 *Isotropic Blobs*

The generation of isotropic blobby patterns relies on Markov/Gibbs random fields (MRF/GRF) [4, 5, 23]. We generate sample images with $K_{ACA} = 4$ classes based on the MRF model described in [43] and [59], which assumes that the only nonzero Gibbs potentials are those that correspond to the one- and two-point cliques. According to this model, neighboring pixels are more likely to belong to the same class than to different classes. The strengths of the two-point clique potentials control the size and shape of the blobs, while the strengths of the one-point clique potentials control the percentage of labels in each class. An iterative procedure is necessary to obtain an MRF image with a given percentage of labels in each class. Finally, we "paint" the resulting sample MRF image with the dominant colors as shown in Fig. 2(e).

For the first empirical study, we generated textures by using the same value for all the two-point clique potentials in order to generate isotropic blobs, which represent the most generic structure that maintains the given color composition.

A.2 *Geometric Blocks*

To generate textures consisting of blocks of different geometries and scales, we propose an iterative algorithm that utilizes randomly placed overlapping fixed sized blocks with the dominant colors. This is a manifestation of the "dead leaves" technique [6, 27, 35].

The iterative block placement procedure is described in Algorithm 1. The upper left corner of each color block is placed at a random position in the image with probabilities given by the desired color percentages. In order to allow random placement of blocks anywhere in the image,
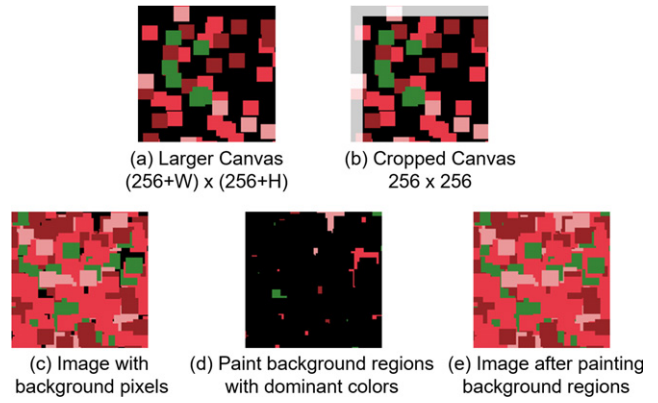

Figure A1. Example of generating geometric blocks. The background canvas is shown in black.

including near the texture borders, we generate a canvas of larger size than the posterized texture image and then crop it back to the right size, as shown in Figure A1. The width of the left and upper strip of the larger canvas equals the size of the color blocks. By including blocks with upper left corners in the strip regions, we enable partial block overlap in the upper and left borders of the cropped image. Note that block overlap at the lower and right borders can happen when the upper left corner of a block is placed near the border.

Algorithm 1: Generate textures with fixed-size color blocks

Input : Dominant colors $\{L_k^*, a_k^*, b_k^*\}$ and percentages p_k for $k = 1, \ldots, K_{ACA}$, thresholds T_h and T_l

Output: Color texture with (approx.) target color percentages

Stage 1: while $D > T_h$ **do**

  Place blocks at random locations with colors $\{L_k^*, a_k^*, b_k^*\}$ selected according probabilities p_k;
  Each block overlaps previously placed blocks;
  Compute the actual color percentages $\hat{p}_k$;
  Calculate the deviation: $D = \sum_k |\hat{p}_k - p_k|$;

end

Stage 2: if *Background pixels exist* **then**

  Extract all connected background regions;
  "Paint" each background region with one dominant color based on the revised color probabilities $\tilde{p}_k$;
  Compute the actual color percentages $\hat{p}_k$;
  Calculate the deviation: $D = \sum_k |\hat{p}_k - p_k|$;

end

Stage 3: while $D > T_l$ **do**

  Place blocks at random locations with colors $\{L_k^*, a_k^*, b_k^*\}$ selected according probabilities $\tilde{p}_k$;
  Compute the actual color percentages $\hat{p}_k$;
  Calculate the deviation: $D = \sum_k |\hat{p}_k - p_k|$;
  Update the probabilities $\tilde{p}_k$;

end

The algorithm consists of three stages. In Stage 1, the probabilities of placing blocks of different colors at random locations of the extended canvas are equal to the desired color percentages. The iterative block placement continues until the deviation D, computed as the sum of the absolute differences between the target and the current color percentages, is below an initial (higher) threshold T_h.

However, as we mentioned above, due to the random block overlap and partial placements (near the borders), the iterative placement procedure is not guaranteed to converge to the desired percentages. So, once the initial threshold is achieved, we need to adjust the placement probabilities in order to facilitate convergence. The iterations will continue in Stage 3, but first we have to make sure that there are no "unpainted" pixels in the canvas.

Fig. A1(c) shows the resulting texture after the first stage. Note that there are some "unpainted" background pixels in the canvas, shown in black in the figure. In Stage 2, we check if there are any "unpainted" pixels in canvas. If this is the case, we use connected component labeling to find all the connected regions in the background. We "paint" each connected region with one dominant color as shown in Fig. A1(d). The fraction of regions painted with each of the dominant colors is based on the revised (see below) color percentages, not taking into consideration the number of pixels in each region; the exact number of pixels assigned to each color at this stage is not important for algorithm convergence.

In Stage 3, the placement probabilities are updated at the end of each iteration in order to enable and accelerate convergence. The idea is to base the probabilities on the differences between the actual and the target color percentages $\Delta p_k = \hat{p}_k - p_k$. The revised color probabilities are then obtained as the normalized differences $\tilde{p}_k = \Delta p_k / \sum_m \Delta p_m$. The third stage ends when the deviation is below a final (lower) threshold T_l.

APPENDIX B. BLOB SCALE ESTIMATION

We propose an efficient method for calculating the average linear scale (in pixels) of the blobs with the same color as shown in Figure B1 and described in Algorithm 2.

Algorithm 2: Estimating the average scale of target color blobs

Input : K-level image I (width W, height H), each level painted with a different (dominant) color, target color (level) C

Output: Estimate of average scale of C-colored blobs

1. Extract all blobs of color (level) C from image I and generate a binary image I_{wat};
2. Apply marker-controlled watershed segmentation to I_{bw} to separate touching blobs;
3. Scan the segmented image I_{wat} from top to bottom, and from left to right, separately;
4. In each scan line, collect all valid segmented lines belonging to the blobs;
5. Calculate the average L_h and L_v of all the horizontal and vertical blob lengths, respectively;
6. Linear scale of target color blobs $S_C = \sqrt{L_h \cdot L_v}$;

Given an input K-level image **I**, each level of which is painted with a different (dominant) color, and a target level (color), we first generate a binary image **I**$_{\mathbf{bw}}$ with only the blobs of the target color **C**. To separate the touching blobs into individual objects, we apply marker-controlled watershed segmentation [46] to the binary image. We first apply the watershed transform to the binary image as a segmentation function. To reduce over-segmented regions generated by the watershed transformation, the segmentation function can be improved by using the marker-controlled watershed segmentation. The marker-controlled watershed segmentation generally utilizes markers for foreground and background to solve the over-segmentation problem. Since the binary image **I**$_{\mathbf{bw}}$ automatically separates foreground
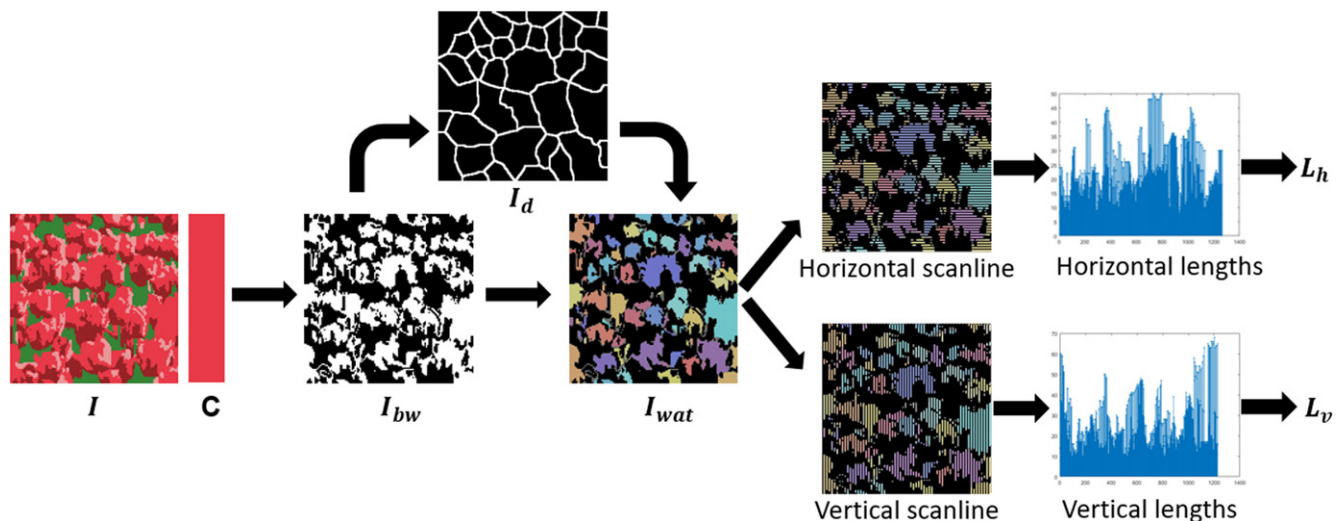

Figure B1. Estimation of the average scale of blobs in given target color (Algorithm 2).

objects (blobs) and background as markers, we can directly fit the segmentation function to its minima at foreground and background locations. $\mathbf{I_d}$ is the ridge line map of the fitted segmentation. $\mathbf{I_{wat}}$ is the segmented result after applying $\mathbf{I_d}$ to $\mathbf{I_{bw}}$. The different colors represent the different connected components.

Once we get the image $\mathbf{I_{wat}}$ with separated blobs, each painted with a different color, as shown in Fig. B1, we apply a line-based calculation to get the average scale. For each row in the image, we collect all segments of horizontal lines belonging to the blobs. The average horizontal length $\mathbf{L_h}$ is the mean of all segments of horizontal lines across the whole image. Similarly, we collect all segments of vertical lines of the blobs and compute the average vertical length $\mathbf{L_v}$. The average area of the target color blobs is the product of $\mathbf{L_h}$ and $\mathbf{L_v}$. We use the square root of the area as the linear scale of the target color blobs.

REFERENCES

[1] B. L. Anderson and J. Kim, "Image statistics for surface reflectance perception," J. Vis. **9**, 1–17 (2009).

[2] J. Beck, "Similarity grouping and peripheral discriminability under uncertainty," Am. J. Psychol. **85**, 1–19 (1972).

[3] J. R. Bergen, "Theories of visual texture perception," in *Spatial Vision, ser. Vision and Visual Dysfunction*, edited by D. Regan (CRC Press, Cambridge, MA, 1991), Vol. 10, pp. 114–134.

[4] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," J. Royal Statist. Soc. B **26**, 192–236 (1974).

[5] J. Besag, "On the statistical analysis of dirty pictures," J. Royal Statist. Soc. B **48**, 259–302 (1986).

[6] C. Bordenave, Y. Gousseau, and F. Roueff, "The dead leaves model: a general tessellation modeling occlusion," Adv. Appl. Probab. **38**, 31–46 (2006).

[7] J. Chen, T. N. Pappas, A. Mojsilovic, and B. E. Rogowitz, "Adaptive perceptual color-texture image segmentation," IEEE Trans. Image Process. **14**, 1524–1536 (2005).

[8] C. Chubb, J. Darcy, M. Landy, J. Econopouly, J. Nam, D. Bindman, and G. Sperling, "The scramble illusion: Texture metamers," in *Oxford Compendium of Visual Illusions*, edited by A. G. Shapiro and D. Todorovic (Oxford University Press, New York, 2015).

[9] Y. Deng, B. S. Manjunath, C. Kenney, M. S. Moore, and H. Shin, "An efficient color representation for image retrieval," IEEE Trans. Image Process. **10**, 140–147 (2001).

[10] M. N. Do and M. Vetterli, "Texture similarity measurement using Kullback-Leibler distance on wavelet subbands," Proc. Int. Conf. Image Proc. **3**, 730–733 (2000).

[11] A. L. Edwards, *Techniques of Attitude Scale Construction* (Ardent Media, New York, NY, 1983).

[12] S. L. Elliott, J. L. Hardy, M. A. Webster, and J. S. Werner, "Aging and blur adaptation," J. Vis. **7**, 1–9 (2007).

[13] R. T. Eskew and R. M. Boynton, "Effects of field area and configuration on chromatic and border discrimination," Vis. Res. **27**, 1835–1844 (1987).

[14] L. He, "A clustering approach for color texture segmentation," Ph.D. dissertation (Northwestern University, Evanston, IL, 2012).

[15] L. He and T. N. Pappas, "An adaptive clustering and chrominance-based merging approach for image segmentation and abstraction," *Proc. Int'l. Conf. Image Proc.* (IEEE, Piscataway, NJ, 2010), pp. 241–244.

[16] Y.-X. Ho, M. S. Landy, and L. T. Maloney, "How direction of illumination affects visually perceived surface roughness," J. Vis. **6**, 634–648 (2006).

[17] Y.-X. Ho, M. S. Landy, and L. T. Maloney, "Conjoint measurement of gloss and surface texture," Psychological Sci. **19**, 196–204 (2008).

[18] Y.-X. Ho, L. T. Maloney, and M. S. Landy, "The effect of viewpoint on perceived visual roughness," J. Vis. **1**, 1–16 (2007).

[19] B. Julesz, "Visual pattern discrimination," IRE Trans. Inf. Theory **8**, 84–92 (1962).

[20] B. Julesz, "Textons, the elements of texture perception and their interactions," Nature **290**, 91–97 (1981).

[21] B. Julesz and B. E. Rogowitz, "AI and early vision – Part II," in *Human Vision, Visual Proc., and Digital Display, Ser. Proc. SPIE*, edited by B. E. Rogowitz (SPIE, Los Angeles, CA, 1989), Vol. 1077, pp. 246–268.

[22] E. Kimura, "Averaging colors of multicolor mosaics," J. Opt. Soc. Am. A **35**, B43–B54 (2018).

[23] R. Kindermann and J. L. Snell, *Markov Random Fields and their Applications* (American Mathematical Society, Providence, RI, 1980).

[24] F. A. A. Kingdom, *Interactions of Color Vision with Other Visual Modalities* (Springer, Cham, 2016), pp. 219–241.

[25] I. Kuriki, "Testing the possibility of average-color perception from multi-colored patterns," Opt. Rev. **11**, 249–257 (2004).

[26] M. S. Landy, "Texture analysis and perception," in *The New Visual Neurosciences*, edited by J. S. Werner and L. M. Chalupa (MIT Press, Cambridge, MA, 2014), pp. 639–652.

[27] A. B. Lee, D. Mumford, and J. Huang, "Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model," Int. J. Comput. Vis. **41**, 35–39 (2001).

[28] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," IEEE Trans. Commun. **COM-28**, 84–95 (1980).

[29] W. Y. Ma, Y. Deng, and B. S. Manjunath, "Tools for texture/color based search of images," in *Human Vision and Electronic Imaging II. Proc. SPIE*, edited by B. E. Rogowitz and T. N. Pappas (SPIE, San Jose, CA, 1997), Vol. 3016, pp. 496–507.

[30] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. Pattern Anal. Mach. Intell. **18**, 837–842 (1996).

[31] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," IEEE Trans. Circuits Syst. Video Technol. **11**, 703–715 (2001).

[32] J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," IEEE Trans. Inform. Theory **IT-20**, 525–536 (1974).

[33] P. Maragos, R. W. Schafer and M. A. Butt (Eds.) *Mathematical Morphology and Its Applications to Image and Signal Processing* (Springer Science & Business Media, 2012), Vol. 5.

[34] P. J. Marlow, J. Kim, and B. L. Anderson, "The perception and misperception of specular surface reflectance," Current Biology **22**, 1909–1913 (2012).

[35] G. Matheron, *Random Sets and Integral Geometry* (John Wiley and Sons, New York, 1975).

[36] J. Maule and A. Franklin, "Accurate rapid averaging of multihue ensembles is due to a limited capacity subsampling mechanism," J. Opt. Soc. Am. A **33**, A22–A29 (2016).

[37] A. Mojsilović, J. Hu, and E. Soljanin, "Extraction of perceptually important colors and similarity measurement for image matching, retrieval, and analysis," IEEE Trans. Image Process. **11**, 1238–1248 (2002).

[38] A. Mojsilović, J. Kovačević, J. Hu, R. J. Safranek, and S. K. Ganapathy, "Matching and retrieval based on the vocabulary and grammar of color patterns," IEEE Trans. Image Process. **1**, 38–54 (2000).

[39] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson, "Image statistics and the perception of surface qualities," Nature **447**, 206–209 (2007).

[40] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Anal. Mach. Intell. **24**, 971–987 (2002).

[41] I. Omer and M. Werman, "Color lines: Image specific color representation," *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)* (IEEE, Piscataway, NJ, 2004), pp. 946–953.

[42] S. Padilla, O. Drbohlav, P. R. Green, A. Spence, and M. J. Chantler, "Perceived roughness of 1/f[beta] noise surfaces," Vis. Res. **48**, 1791–1797 (2008).

[43] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," IEEE Trans. Signal Process. **SP-40**, 901–914 (1992).

[44] T. N. Pappas, J. Chen, and D. Depalov, "Perceptually based techniques for image segmentation and semantic classification," IEEE Commun. Mag. **45**, 44–51 (2007).

[45] T. N. Pappas, D. L. Neuhoff, H. de Ridder, and J. Zujovic, "Image analysis: Focus on texture similarity," Proc. IEEE **101**, 2044–2057 (2013).

46 K. Parvati, P. Rao, and M. M. Das, "Image segmentation using gray-scale morphology and marker-controlled watershed transformation," Discrete Dyn. Nat. Soc. **2008** (2009).

47 B. Pinna and S. Grossberg, "The watercolor illusion and neon color spreading: a unified analysis of new cases and neural mechanisms," J. Opt. Soc. Am. A **22**, 2207–2221 (2005).

48 A. B. Poirson and B. A. Wandell, "Appearance of colored patterns: pattern-color separability," J. Opt. Soc. Am. A **10**, 2458–2470 (1993).

49 Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," Int. J. Comput. Vis. **40**, 99–121 (2000).

50 H. S. Sawhney and J. L. Hafner, "Efficient color histogram indexing," *Proc. Int'l. Conf. Image Proc.* (IEEE Computer Society, Washington, DC, 1994), pp. 66–70.

51 N. C. Silver and W. P. Dunlap, "Averaging correlation coefficients: Should Fisher's z transformation be used?," J. Appl. Psychology **72**, 146 (1987).

52 M. Swain and D. Ballard, "Color indexing," Int. J. Computer Vision **7**, 11–32 (1991).

53 E. Switkes, A. Bradley, and K. K. D. Valois, "Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings," J. Opt. Soc. Am. A **5**, 1149–1162 (1988).

54 L. L. Thurstone, "A law of comparative judgment," Psychological Review **34**, 273 (1927).

55 J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles* (Addison, Reading, MA, 1974).

56 A. J. van Doorn, H. de Ridder, and J. J. Koenderink, "Pictorial relief for equiluminant images," in *Human Vision and Electronic Imaging X, Ser. Proc. SPIE*, edited by B. E. Rogowitz, T. N. Pappas, and S. J. Daly (SPIE, San Jose, CA, 2005), Vol. 5666.

57 J. Webster, P. Kay, and M. A. Webster, "Perceiving the average hue of color arrays," J. Opt. Soc. Am. A **31**, A283–A292 (2014).

58 M. W. A. Wijntjes and S. C. Pont, "Illusory gloss on Lambertian surfaces," J. Vis. **10**, 1–12 (2010).

59 J. Zujovic, "Perceptual texture similarity metrics," Ph.D. dissertation (Northwestern University, Evanston, IL, 2011).

60 J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Structural similarity metrics for texture analysis and retrieval," *Proc. Int'l. Conf. Image Proc.* (IEEE, Piscataway, NJ, 2009), pp. 2225–2228.

61 J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Structural texture similarity metrics for image analysis and retrieval," IEEE Trans. Image Process. **22**, 2545–2558 (2013).