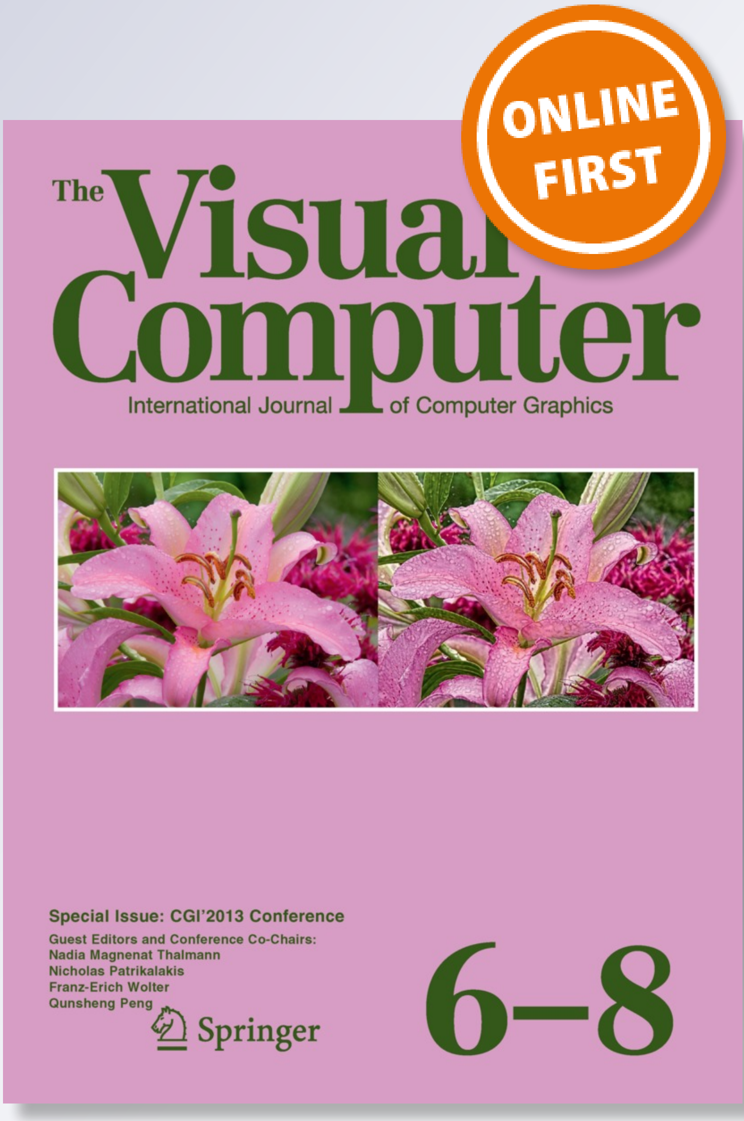
Grey conversion via perceived-contrast

Wei Zhu, Ruizhen Hu & Ligang Liu

The Visual Computer
International Journal of Computer Graphics

ISSN 0178-2789

Vis Comput
DOI 10.1007/s00371-013-0854-9



Springer

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Springer

Author's personal copy

Vis Comput
DOI 10.1007/s00371-013-0854-9

ORIGINAL ARTICLE

Grey conversion via perceived-contrast

Wei Zhu · Ruizhen Hu · Ligang Liu

© Springer-Verlag Berlin Heidelberg 2013

Abstract This paper presents a new color-to-gray conversion algorithm capturing the perceived appearance of color images. Based on the *Filter Theory*, we formulate a novel measurement of channel-level distinction, called Channel Salience, to depict the filter level of three color stimuli. This salience metric guides a contrast adjustment process to enhance the perceived grayscale in the final output with a two-steps conversion. Experimental results show that our algorithm produces pleasing results for a variety of color images and we further extend the Channel Salience to edge detection.

Keywords Color-to-gray conversion · Filter theory · Channel salience · Perceived contrast

1 Introduction

There is a strong demand for color-to-gray conversion in many aspects, e.g., it helps us economically print color images via monochrome version; it enables the application of single channel algorithms on color images, like Canny operator [8] for edge detection; photographers take color images and then remove color information to draw viewers' attention on outstanding contents. In a sale of important nineteenth and twentieth century photographs by New York's Swann Galleries in 2008, 364 out of 389 were monochrome photographs [12], which shows the high popularity of black and white photos. All these applications prompt researchers to develop better color-to-gray conversion algorithms.

In the last decade, many approaches have been proposed, mainly focusing on preserving the color distance and reducing the information loss [13, 14, 16, 29]. Methods proposed in [13, 14, 16] consider global contrast in different channels. When the contrast in channels are not collaborative, the final output would be muddy. For local detail preserving method [29], details could be overenhanced and the results hardly reflect global perceptual impact. We present a synthetic color image (upper middle), with contradicting patterns in difference color channels, to demonstrate that including all the information does not always lead to a better result. Most people would feel this color image smoothly gets lighter and then darker from left to right, and it is darker on the left side. Our result fits this trend better because we does not pick up every bit of all information, but concentrates to the perceived contrast. Figure 1 (lower row) presents the results of 4 state-of-the-art algorithms.

To better capture the perceived grayscale, we come up with a new algorithm based on the *Filter Theory* proposed by Broadbent [7]. Broadbent studied the filter property of human perception and modeled it as selective processing. Later experiments conducted by Treisman [31] showed that human brain processes stimuli with bigger distinction faster and nonattended information passes the processing bottleneck in an weakened fashion. In this paper, we consider each channel as a stimulus and measure its level of suppression by exploiting a new feature, called *Channel Salience*. Then contrast adjustment is proposed to enhance contrast of the

Electronic supplementary material The online version of this article (doi:10.1007/s00371-013-0854-9) contains supplementary material, which is available to authorized users.

W. Zhu (✉) · R. Hu
Department of Mathematics, Zhejiang University, Hangzhou 310027, China
e-mail: weizhu.zju@gmail.com

L. Liu
School of Mathematical Sciences, University of Science and Technology of China, Anhui 230026, China
e-mail: ligang.liu@gmail.com

Published online: 29 May 2013

Author's personal copy

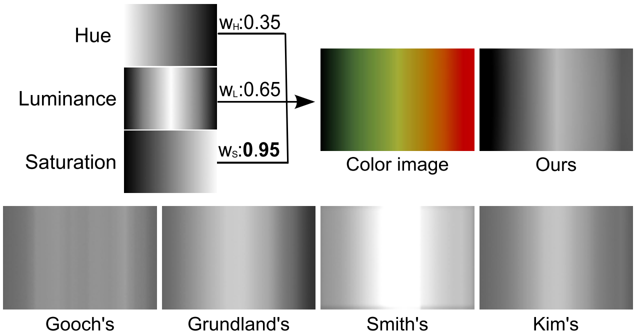


Fig. 1 Gray conversion results on a combined color image (*upper middle*) with mixture weights (0.35, 0.65, 0.95). Comparison with 4 state-of-the-art algorithms Gooth et al. [13], Grundland et al. [14], Smith et al. [29], and Kim et al. [16] (*lower*), our result could capture perceived contrast in saturation channel

most salient channel in global manner, so that final gray results would be more similar to what people perceive. Our result in Fig. 1 (upper right) better captures the perceived grayscale, which reflects the saturation changes in the color image.

2 Related work

Color-to-gray conversion Conventional gray conversion algorithms proposed in the last decade mainly fell into two categories, local adjustment and global adjustment. On one hand, local adjustment approaches manipulated pixel intensity according to local color difference. These methods combined high-frequency chromatic information with luminance and preserved local chromatic edges but sacrificed global consistence [6, 29]. Based on the experiments of the Coloroid system, Neumann et al. [21] formed consistent gradient and direct integration to get gray transformation. However, these algorithms tried to visualize all details, which would bring artifacts.

On the other hand, global adjustment approaches tried to minimize difference between distances defined on entire color image and on gray image. Approaches captured the color distance between all pairs of pixels [13] or all pairs of colors [25], and they built a quadratic energy function to solve the gray images. With a similar energy function as in [13], Kim et al. [16] proposed a nonlinear global mapping model, which could achieve the mapping consistency. Grundland and Dodgson [14] proposed the Decolorize algorithm that globally convert the color to grayscale with a continuous, image-dependent mapping. Using mass spring model, another color-to-gray mapping [17] iteratively added the chromatic difference to luminance. Kuk et al. [18] formed a new energy term to balance the gradient among pixels, or between pixels and some predetermined landmarks. These landmarks were obtained by color quantization. Zhao and Tamimi [34] performed color-to-gray conversion in the spectral domain and generated results with an inverse transformation. Cui et al. [9] introduced ISOMAP to form a manifold reduction method between color space and gray space. Lu et al. [20] reduced strict order constrain in the previous method and maximally preserved the original color contrast. However, most global algorithms paid attention to the preservation of color distance and do not consider perceived contrast of color representation.

Moreover, some researchers considered the grayscale problem, not only as final output, but also as preoperation of other algorithms. Atcuti et al. built two salience catching algorithms, one produces stable SIFT matching results [2] in which the saliency is computed per entire image. Another preserves color salience map consistent results [3] in which saliency is computed for each (R, G, B) color channel. Another algorithm based on image fusion was described in [4].

As far as we know, all of these algorithms evenly treated every channel of color image and combined the difference together to preserve the whole information. Nevertheless, color-to-gray conversion inevitably loses color information and it is not possible to protect all channels equally.

Evaluation scheme Čadík [10] introduced a psychological evaluation of 7 color-to-gray algorithms. There were two experiments: accuracy and preference. Apparent grayscale [29] ranked the best in accuracy experiment and Decolorize [14] ranked the best in preference experiment. Neither of them outperformed each other. Specifically, each of the 7 conversion schemes was ranked the worst for at least one image [10]. In this comprehensive evaluation, the absence of universal satisfactory result showed lack of perceived grayscale interpretation and drove development of new gray conversion.

Color representation Pridmore et al. [22] suggested that perceived brightness is linked to three digital color channels: hue, luminance, and saturation. And these three channels were concluded to be sufficient to describe any perceived color [33]. One advantage of such experimental setting is that these attributes are generally understood [27]. To make our conversion tightly cooperate with phycological experiments, we use luminance, saturation, and hue related color space. Quite a few color spaces are available: HLS, HIS, HSV, CIE LCHab, and Improved HLS (IHLS), while saturation in IHLS color space [15] was normalized and proved to be independent of the luminance. IHLS color space forms a more independent subsystem which would give a more stable foundation for our perceived contrast analysis and it has been widely used in many applications [5, 11, 23, 26]. We adopt IHLS color space as color representation and follow the original definition in [15]. Color images are transformed into three channels, i.e., hue (H), luminance (L), and saturation (S).

Author's personal copy

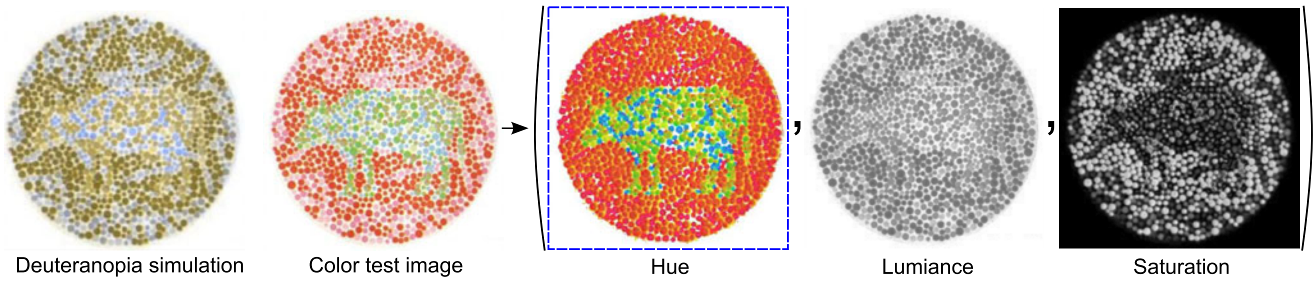


Fig. 2 Color blind test image, color simulation of deuteranopia deficiency and hue, luminance, and saturation channel in IHLS color space. The perceived important information is encoded in hue channel shown in *blue dashed rectangle*. Channel Salience defined in this paper is H:0.604, L:0.070, and S:0.377, respectively

3 Channel salience

The *Filter Theory* proposed by Broadbent [7] suggests that human perception is a selective processing and it has a filter to prevent the information processing system from overloading. Concluded by this theory, not full stimuli can be perceived and only part of them is "accepted" while others are suppressed. In [31], one of experiments gives a further clue that information filtered or not highly depends on the distinction of that stimulus. The color-to-gray conversion could be considered as a selective process and the perceived appearance is more related to the most distinctive stimulus.

Three conceptual stimuli: wavelength, purity, and luminance are utilized in many psychological experiments. Noted by Pridmore [22], these three stimuli link to three color channels: hue, saturation, and luminance, which could affect the perceived appearance. In Fig. 2, we present an example of color blindness test. For people with deuteranopia (green and purplish-red blindness) deficiency, only a deer in the saturation channel can be recognized, while people without deficiency would see the clear cow encoded in hue channel and hardly recognize the ambiguous deer in saturation channel. Based on this observation, we adopt the IHLS color space and assume each independent channel as one stimulus when considering the Filter Theory in color-to-gray conversion.

The next step is to define Channel Salience, the distinction score of each stimulus, which measures perceived strength of each channel. There are few experiments discussed perceived importance of channel stimuli before. One heuristic idea of distinction measurement is borrowed from black-and-white photograph. Photographers capture color pictures and convert them into black-and-white professionally. Photo subjects are usually emphasized by presenting high contrasts between two polarities between black and white [12]. They pay much attention on highlights and shadows; see Fig. 3. The luminance contrast between bright face and dark shadow, or the hue contrast between yellow and green field, composes main contrast which is well elaborated in gray images created by experts. Note that most regions in the two images are highly contrast and the areas of two extremes are having similar areas. Thus, channel salience could be defined as the contrast between two polarities in this channel, and salient channel with high distinction should have two well-separated groups of entries with similar sizes.

Fig. 3 *Left*: the original color images; *Middle*: the black-and-white images converted by experts; *Right*: our gray conversion results

We define Channel Salience with two criteria to measure salience of each channel:

- The channel with higher salience should have a clear cut of two separating groups of entries.
- The number of entries in two groups should be relatively comparable.

We formulate salience model which fits these requirements with two functional terms:

$$CS(ch) = CS_{\mathrm{D}}(ch) * CS_{\mathrm{B}}(ch) \quad (1)$$

where *ch* is one of the three channels in IHLS color space. The first term CS_{D} is the distinction term, measurement of distinction of two entries; and the second term CS_{B} is the balance term, measurement of quantitative balance of two entries.

Author's personal copy

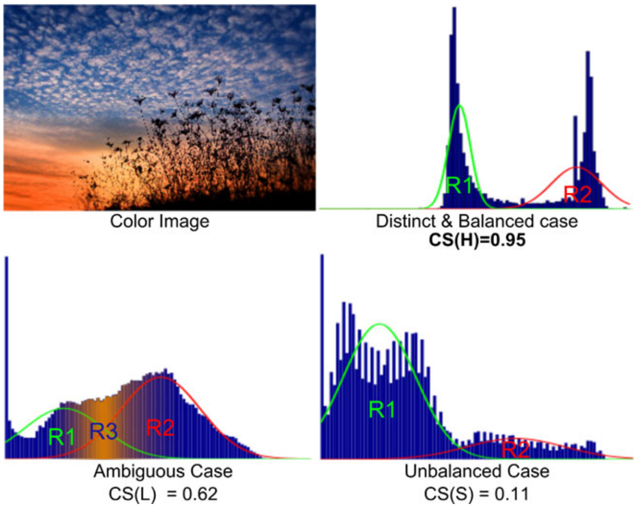


Fig. 4 Channel Salience computation of hue (*upper right*), luminance (*lower left*) and saturation (*lower right*) channel of the color image (*upper left*). Both distinct and balanced channel (hue) has the highest salience value

Distinction term In order to analyze polarities of each channel, a general Gaussian Mixture Model [28] with two centers is applied to generate a probability model of pixels' distribution. We get two posterior probability density functions $P_{ch}(x|c_1)$, $P_{ch}(x|c_2)$, where $P_{ch}(c_1)$, $P_{ch}(c_2)$ are prior probabilities and $\mu_{ch}(c_1)$, $\mu_{ch}(c_2)$ are mean. More technical details would be discussed later.

We assume that salient information composes of two separated entries, thus pixels can be efficiently classified into two clusters. A histogram-based method is adopted to measure the classification efficiency. Hu et al. [24] defined the classification efficiency by dividing the classified possibility into three parts: two consistent regions R_1, R_2 and an inconsistent region R_3 (see Fig. 4 Ambiguous Case). Two probability density functions are close to each other in R_3, thus data in R_3 has ambiguous classification, which violates our first requirement of the data distinction. Hence, in order to punish data points lying in ambiguous region, we define distinction term as the following:

$$CS_{\mathrm{D}}(ch) = \left(1 - \frac{1}{N}\sum_i A_i * n_i\right) * (1 - C)$$

where $A_i = \exp(-(P_{ch}(x_i, c_1) - P_{ch}(x_i, c_2))^2 / 2\sigma_{\mathrm{post}}^2)$ measures ambiguity of ith bin with x_i as its index, $P_{ch}(x_i, c_k) = P_{ch}(c_k)P_{ch}(x_i|c_k)$ $(k = 1, 2)$ is the joint probability of ith bin belonging to cluster k, σ_{post} adjusts the tolerance of ambiguity and we set $\sigma_{\mathrm{post}} = 0.5$ in our experiment, n_i is the number of pixels fallen into ith bin, and N is the total number of pixels. $C = \exp(-(\mu_{ch}(c_1) - \mu_{ch}(c_2))^2 / 2\sigma_{ch}^2)$ adopts the distance between two mean values to compute the contrast of two clusters. The setting of parameter σ_{ch} will be discussed later in this section.

An ambiguous histogram is shown in Fig. 4 (lower left) (bins with high ambiguity is colored in orange and the transparency reflects the magnitude of ambiguity). The ambiguity is high if two probability density functions are close (R_3), on the contrary, it is low if they are far from each other (R_1 and R_2). From the color image, we can see that the average distributed luminance channel is not perceived most strongly. As a result, the luminance channel scores a lower channel salience. Instead, large amount of red and blue pixels in hue channel constitute a promising big contrast. A highly distinctive case (Fig. 4, upper right) shows two non-overlapping clusters, which promises a dominating contrast in hue channel.

Balance term The number of data points in two clusters has to be comparable. Otherwise, the smaller cluster would not be able to support a highly perceived distinction. Thus, we give the definition of balance term as follows:

$$CS_{\mathrm{B}}(ch) = \exp\left(-\left(P_{ch}(c_1) - P_{ch}(c_2)\right)^2 / 2\sigma_{\mathrm{prior}}^2\right)$$

where σ_{prior} adjusts the tolerance of difference in cluster size (usually set as 0.5 in our experiments).

An unbalanced case is shown in Fig. 4 (lower right). Two components may have a high classification efficiency but may differ a lot in size. If two prior probabilities (the areas of R_1 and R_2) are different, the tiny part would not attract enough attention, which gives a low salience. From the color image (upper left), we can see that only a small part is highly saturated and more regions are dull, hence the contrast in saturation channel is not strongly perceived. Balanced case in Fig. 4 (upper right) shows that the distribution of hue forms two clusters with comparable size which promises a eye-catching distinction in hue channel.

Technical details In IHLS color space, we model distributions of pixel values, without spatial considerations, of three channels separately using Gaussian mixture (GMM) with two Gaussians, which measures the possibility and distinction of two clusters in each channel. However, if channels are scattered of a distribution with three or more clusters, in this case, we group those data close to each other into one cluster and compute the channel salience with two grouped clusters. With the benefit of GMM, we can minimize the grouping errors and get reasonable clusters; please see luminance histogram in Fig. 4 and hue histogram of in Fig. 6. We adopt the modeling methodology and maximum likelihood solver in [28] to assign every pixel likelihoods of belonging to each component. If solver could not converge, we simply set the channel salience as zero. And in the hue channel, we only include pixels with luminance and saturation between 0.2 and 0.8.

After GMM is found, we discuss the parameter setting for contrast weight C. In the three channel histograms of

Author's personal copy

Fig. 5 Results of different perceived contrasts. Contrasts are computed by three channels respectively and combination of all channels via Eq. (3). The salient channel promises a result more close to perceived appearance

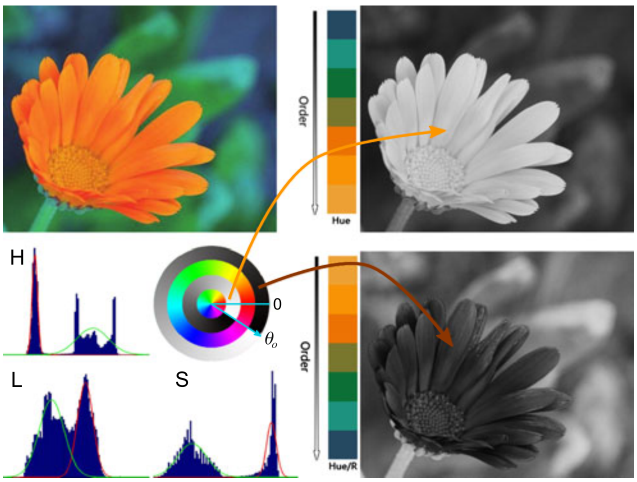


Fig. 6 The bipolarity situation in hue channel. After getting the optimized angle θ_o, we could assign the hue order as counterclockwise (our experiment setting) or clockwise

Fig. 6, ambiguity measurements A_i are both very small in hue and saturation channel and balance term is similar, also. However, in color perception experiment [35], hue channel gives us larger perceived contrast than saturation channel. The contrast weight C endeavors to adjust this difference. Suggested by the experiment result [35], we set $\sigma_H = 0.2$ and $\sigma_{L,S} = 0.4$.

It should be noticed that the hue channel is cyclic, raising a toric topology problem. An optimization is performed to maximize the distinction term by cutting the hue circle into a linear axis among some pre-defined angles $\frac{\pi}{6}i, i = 0, 1, \ldots, 11$. The optimized angle not only solves the toric topology problem, but also captures the maximal perceived hue contrast. Considering polarity, we set yellow and red colors brighter than blue and green colors (similar experiment setting as Wyszecki [32]). An example is given to illustrate a bipolarity situation (Fig. 6). The default setting gives a brighter flower, on the contrary, the reversed setting gives a darkened flower, both of these two results are reasonable and allowed.

4 Color-to-gray conversion

We now describe main framework of color-to-gray conversion cooperating with Channel Salience defined in previous section. The conversion is modeled as a global contrast adjustment on luminance channel L. The perceived brightness is mainly determined by luminance and could also be affected by saturation and hue, which has been adopted by previous works [3, 14, 16, 17, 29]. We build a contrast map M to adjust L and obtain the final conversion result G:

$$G = L + \alpha M \tag{2}$$

where α constrains the outlying pixels ($G \leq 0$ and $G \geq 1$) less than 5 %. The contrast map M, which stores the salient color contrast, can be reconstructed from the perceived contrast ΔM.

4.1 Perceived contrast

With Channel Salience, one straightforward way to model the perceived contrast is to define it as a weighted sum of three channels' contrast:

$$\Delta M = \omega_H \Delta H + \omega_L \Delta L + \omega_S \Delta S \tag{3}$$

where ω_{ch} ($ch = H, L, S$) is the normalized salience of each channel computed via Eq. (1), $\sum \omega_{ch} = 1$. And ΔH, ΔL, ΔS are the difference of three normalized channels. But a weighted sum does not always give a good result. When the channels are not cooperative, the result would be muddy. Results obtained from the preliminary experiments also show that the weight sum would offset the information from different channels, making the results muddier and lost the salient contrast. One example is shown in Fig. 5. We can see that the result with weighted contrast (right most) is muddy and gives low impact.

Instead, we select one single channel with the highest Channel Salience, i.e., salient channel ch^s. This channel forms more important contrast we perceived in the original color image (see Fig. 2). It is also because if any other channels are having a similar contrast, there is no need to include them into the perceived contrast. On the contrary, if they are contradictive, adding contradictive channels together would

Author's personal copy

make the results muddy (Figs. 1 and 5). Figure 5 shows results with different channels as the salient channel. From the color sequences of different channels, we can see that the salient channel (hue) discriminates sunflowers (yellow) and lawn (green) clearly, which reflects salient contrast that should be maintained in gray images. Hence, our perceived contrast is defined as follows:

$$\Delta M = \Delta ch^s \tag{4}$$

where $\Delta ch^s_{ij} = ch^s_i - ch^s_j$ is the difference between two pixels i and j in the normalized salience channel, ch should be the salient channel among H, L, or S.

4.2 Adjustment

Similar to other algorithms [6, 16, 17, 21, 29], we add the perceived contrast onto the luminance channel. We construct the final gray image from the sum of the luminance and perceived contrast, then minimize the following energy function:

$$\min_g \sum_{i\in\Omega}\sum_{j\in\Omega}\left(g_i - g_j - (\Delta L_{ij} + \Delta M_{ij})\right)^2 \tag{5}$$

where Ω is the set of all pixels. Here, we could achieve a quicker solution using the method in [30]. Tanaka et al. [30] proved that if using a conjugate gradient solver and giving luminance channel as initialization, solution of Eq. (5) at ith pixel would be $g_i = L_i + \frac{1}{N}\sum_{j=1}^{N}\Delta M_{ij}$ and N is the total number of pixels. and we can build the contrast map M in Eq. (2):

$$M_i = \frac{1}{N}\sum_{j\in\Omega}\Delta M_{ij}, \quad \forall i \in \Omega \tag{6}$$

Optimization of Eq. (5) would be vulnerable to large quantity of noise. And from Eq. (6), the adjustments on these pixels are highly related to values in salient channel, i.e., gray values are changed equally if values of these pixels in salient channel are same. This would cause artifacts; see Fig. 7 for an example. The reason is that the type of adjustment only considers the difference between pixel values but the spatial information of pixels is neglected, the grayscale is not smooth. Especially when the salient channel is a hue channel, pixel values are nearly piecewise constant (upper right) which would bring unwanted sharp edges in the final gray image (lower right). In order to reduce these artifacts, we separate the conversion process into following two steps: tone elaboration and propagation.

Tone elaboration We first perform a nonuniform sampling to select the same number of representative points from each cluster of GMM according to the Gaussian distribution of the salient channel ch^s built in Sect. 3. Only 1 ‰ of the total number of pixels is selected. Figure 7 shows the sampling possibility (middle right) and the sampled representative points of a color image (middle left). Points with two different colors belong to two clusters. We change set of all pixels Ω in Eq. (6) to set of representative pixels Ω_R,

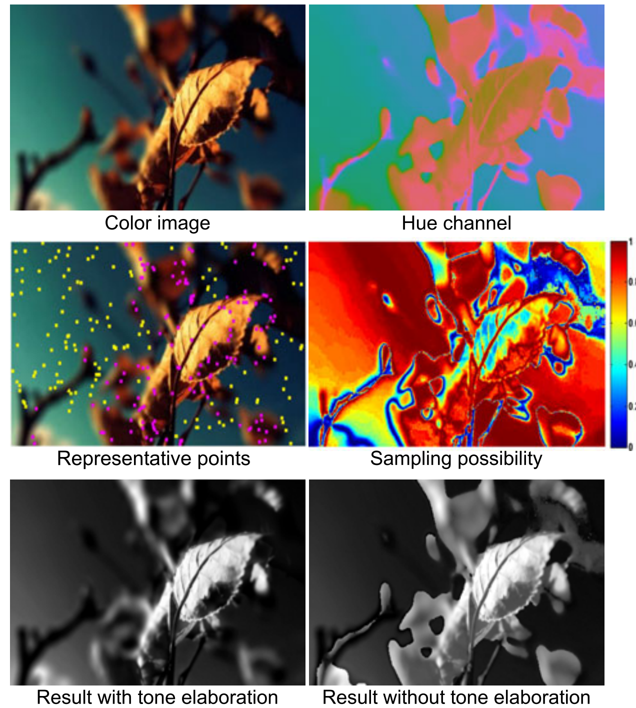


Fig. 7 Color-to-gray conversion. The result with tone elaboration (*lower left*) first enhance the perceived contrast more smooth than directly constructing from all pixel pairs (*lower right*)

$$E_k = \frac{1}{N}\sum_{j\in\Omega_R}\Delta M_{kj}, \quad \forall k \in \Omega_R \tag{7}$$

Elaborated tone E enforces contrast map of sampling points to match the salient contrast and enlarges distance of sampling pixels between different clusters.

Propagation With the elaborated tone E, we propagate them to all pixels via the method applied in [1] which considers both color and spatial information. The propagation energy function is defined as follows:

$$\min_M \sum_{i\in\Omega}\sum_{k\in\Omega_R} z_{ik}(M_i - E_k)^2 + \lambda \sum_{i\in\Omega}\sum_{j\in\Omega} z_{ij}(M_i - M_j)^2$$

where weight $z_{ij} = \exp(-\Delta D_{ij}{}^2/\sigma_d)\exp(-\Delta M_{ij}{}^2/\sigma_m)$, ΔD_{ij} and ΔM_{ij} are the spatial distance and value distance of salient channel between pixel i and j, σ_d and σ_m are set as 100 and 0.1, furthermore the parameter $\lambda = 0.5$ controls the smoothness of the propagation result. After getting the entire contrast map M, we add luminance channel and obtain our final gray conversion results using Eq. (2).

Figure 7 (lower left) shows one example of propagation result, we can see that propagation result could catches the

Author's personal copy

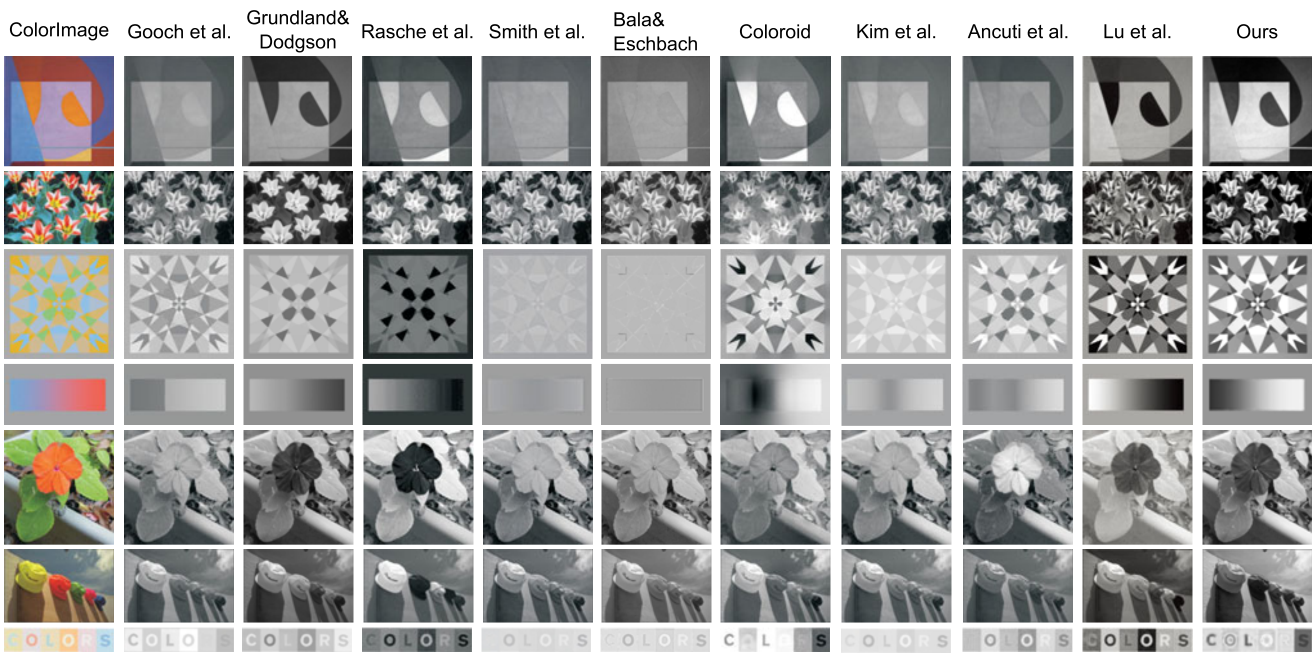


Fig. 8 Comparison with state-of-the-art algorithms: Gooch et al. [13], Grundland and Dodgson [14], Rasche et al. [25], Smith et al. [29], Bala and Eschbach [6], Neumann et al. [21], Kim et al. [16], Ancuti et al. [3], Lu et al. [20], and our results

contrast of salient channel precisely and ensures a pleasing result while direct reconstruction from all pixel pairs using Eq. (6) fails (lower right).

5 Results and discussion

Our framework comprises two major parts, the computation of GMM and gray conversion by propagation. For an image of 1024 × 768, the GMM takes 2.9 seconds in average and the gray conversion takes 3.4 seconds in average in MATLAB of 32-bit PC with dual core 2.93 GHz CPU and 4 GB ram. Results are shown in Figs. 8, 11.

Saliency-guided decolorization [3] and ours both consider perceived salience. Ancuti et al. developed decolorization which put emphasis on maintain salience map before and after the conversion. Their algorithm computed saliency maps for each (R, G, B) color channel. However, our method does not adopt the salience map but builds channel salience, which is original for contrast comparison with three channels (hue, luminance, saturation) and it could help to preserve the most salient information in perceived grayscale.

And another novel algorithm [20] reduced strict order constrain in the previous method and maximally preserved the original color contrast. As well as other methods, it treated all contrast equally and did not consider salient perceived contrast. Our method not only considers the perceived contrast, but also kindly picks color order from the salient channel; see Figs. 4, 6. We compute color contrast

Table 1 Color contrast preserving ratio (CCPR) comparison with Smith et al. [29], Gooch et al. [13], Kim et al. [16], and Lu et al. [20]

τ	[28]	[12]	[15]	[19]	Ours
1	0.70	0.69	0.72	0.76	0.70
5	0.61	0.63	0.64	0.72	0.67
10	0.55	0.55	0.56	0.66	0.62
15	0.51	0.5	0.5	0.61	0.60

preserving ratio (CCPR) in [20] for comparison in Table 1. It should be noticed that this paper enhances main contrast in the salient channel, not all contrast contained in color images.

5.1 User study

To compare with previous methods, we have conducted a user study. There were totally 25 various color images, of which 5 were the same as those used in [10] and others were images of varying characteristics, including city, landscape or portrait selected from Flickr. Please refer to the supplementary material for the complete set of results.

We invited 45 participants (15 males) at the age of 18 to 45. Among them, 10 were professional photographers. All the participants have normal or corrected-to-normal eyesight, no color deficiency were reported and none of them have done any computer graphics or related work. The images were displayed on a calibrated 27″ monitor with IPS panel and experimental venue was well illuminated with

Author's personal copy

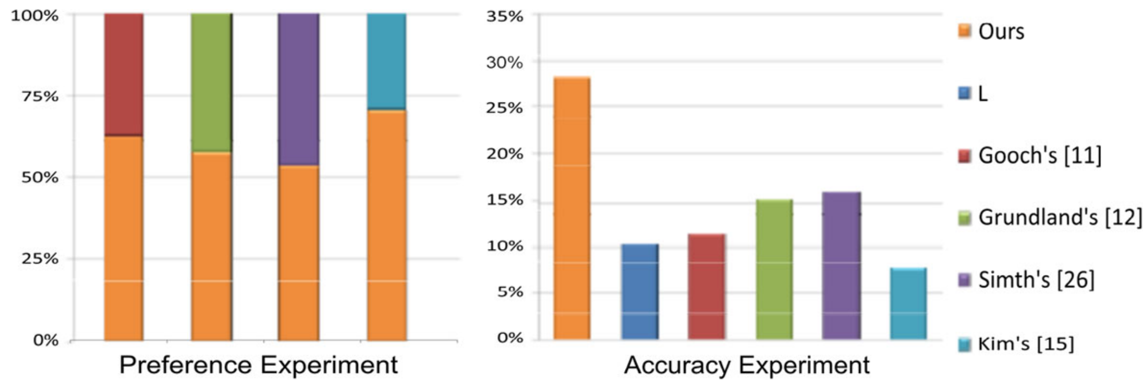


Fig. 9 Results of user study with 4 state-of-the-art algorithms

white fluorescent lamp (4100 K), then we asked them to take two experiments.

Preference experiment Each time, two gray images were displayed on the screen, one of which was ours while the other was randomly chosen from Gooch et al. [13], Grundland et al. [14], Smith et al. [29], and Kim et al. [16]. Participants were instructed to select the image they preferred. Note that the preference experiment was carried first and the participants had never looked at the color version of these images before. Each participant was subjected to 10 questions in this experiment.

Accuracy experiment Each time, the original color image and 6 gray images (our result, luminance channel and 4 other methods Gooch et al. [13], Grundland et al. [14], Smith et al. [29], and Kim et al. [16]) were presented to the participants in a random order. Participants were asked to select the one that best represents the color version. Each participant was subjected to 15 questions in this experiment.

The results of user study are shown in Fig. 9. For preference experiments (left), we present the proportion of participants who preferred our results over those results generated by the other 4 algorithms. The participants generally give a higher preference to the results from our algorithm. For accuracy experiments (right), we present the distribution of choice for each algorithms which are selected by participants as the best representing. We can see that our approach also achieved better performance compared to other algorithms, which reflects that our algorithm is more close to human perception.

The algorithm of Smith et al. [29] gets comparable score to us in preference experiments, while sometimes it produces local artifacts since the detail enhancement would increase unwanted edges. However, our propagation would constrain similarity of neighborhood and this prevents local artifacts. In algorithms [13, 14, 16], they do not consider the salient stimulus and would provide flat results for many images (see Fig. 11). In this paper, we compute Channel Salience to find the salient channel, thus the results better capture the perceived contrast in a global manner.

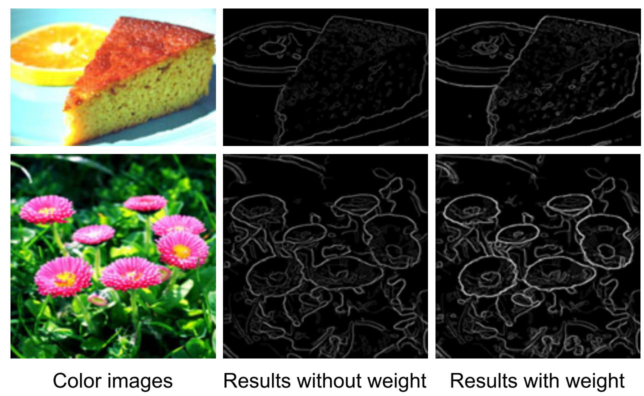


Fig. 10 Application: edge detection with Channel Salience. Our results (*right*) better capture the salient edges

5.2 Application: edge detection

Here we propose an application based on Channel Salience defined in Sect. 3. Edge detection in color image is a fundamental problem in image processing and computer vision. Edges are usually found by computing the local gradient. For color images, Lee and Cok [19] used one local vector gradient to extend the Canny's edge detector [8] to locate the boundaries. However, perceived edges are perceptual results following the *Filter Theory*. Hence, local contrast of one channel should be highlighted if it is salient and suppressed if it is not. In this way, Channel Salience in this paper could contribute to a perceived gradient. With edge detector [19], we calculate the vector gradient in IHLS color space weighted by their normalized Channel Salience, and then carry out the extended Canny detector to identify the edges. This simple modification would increase the contribution of channels with high contrast and reduce the disturb of evenly distributed channels. Figure 10 shows two results of the edge detector used in [19] with and without our Channel Salience related weight.

Author's personal copy

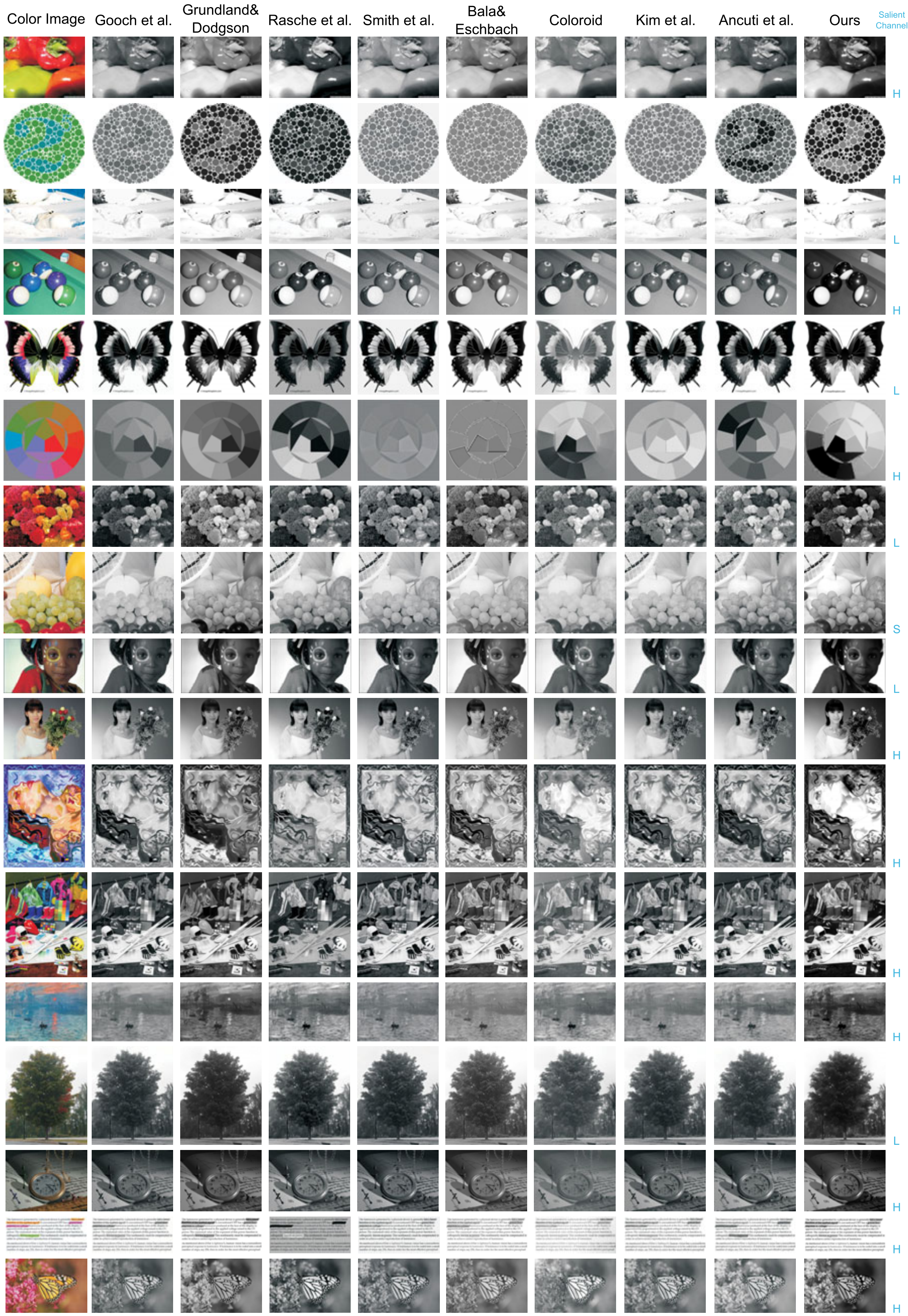


Fig. 11 Comparison with state-of-the-art algorithms: Gooch et al. [13], Grundland and Dodgson [14], Rasche et al. [25], Smith et al. [29], Bala and Eschbach [6], Neumann et al. [21], Kim et al. [16], Ancuti et al. [3], and our results. The last column marks the salient channel of each image

Author's personal copy

6 Conclusion and future work

Color-to-gray conversion is a complex phycological problem, thus the perceived contrast should be considered. In this paper, we present a perceived color-to-gray conversion algorithm, which comprehensively reproduces the perceived appearance of a color image in grayscale version. Based on the *Filter Theory*, we introduce the concept of Channel Salience to measure the contrast level of different channels. Instead of combining information from all channels, we pick the channel with highest salience and enhance its contrast in the conversion process. To make the final gray image more smooth and maintain the contrast in a more efficient and emphatic way, we separate the conversion into two steps: tone elaboration and propagation.

Experimental results show that our algorithm produces pleasing results for a variety of color images. Extending image decolorization to video is presented in many papers. To achieve temporal coherence, correspondence between frames has to be found. We can include coherent information as constraints in the edit-propagation process and extend our framework to video. In some extreme cases, the computation of GMM is not stable. Under this situation, we should carefully pick the initial parameters. Furthermore, when two channels, even three, are almost equally strong, it is difficult to tell how information is processed by human brain and hard to decide which channel takes over others. It happens when color images contain too much information. In this dilemma situation, we should consider more measurements, like fitness of GMM components.

References

1. An, X., Pellacini, F.: AppProp: all-pairs appearance-space edit propagation (2008)
2. Ancuti, C.O., Ancuti, C., Bekaert, P.: Decolorizing images for robust matching. In: 2010 IEEE International Conference on Image Processing, pp. 149–152 (2010)
3. Ancuti, C.O., Ancuti, C., Bekaert, P.: Enhancing by saliency-guided decolorization. In: Proceedings of the IEEE Computer Vision and Pattern Recognition, CVPR'11 (2011)
4. Ancuti, C.O., Ancuti, C., Hermans, C., Bekaert, P.: Image and video decolorization by fusion. In: Computer Vision—ACCV 2010, vol. 6492, pp. 79–92 (2011)
5. Angulo, J., Serra, J.: Color segmentation by ordered mergings. In: 2003 IEEE International Conference on Image Processing, vol. 2, pp. 125–128 (2003)
6. Bala, R., Eschbach, R.: Spatial color-to-grayscale transform preserving chrominance edge information. In: 12th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications, pp. 82–86 (2004)
7. Broadbent, D.: Perception and Communication. Pergamon Press, Oxford (1958)
8. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6), 679–698 (1986)
9. Cui, M., Hu, J.X., Razdan, A., Wonka, P.: Color-to-gray conversion using ISOMAP. Vis. Comput. **26**(11), 1349–1360 (2010)
10. Čadík, M.: Perceptual evaluation of color-to-grayscale image conversions. Comput. Graph. Forum **27**(7), 1745–1754 (2008)
11. Fleyeh, H.: Color detection and segmentation for road and traffic signs. In: 2004 IEEE Conference on Cybernetics and Intelligent Systems, vol. 2, pp. 809–814 (2004)
12. Freeman, M.: The Complete Guide to Black & White Digital Photography. Lark Books, Asheville (2009)
13. Gooch, A.A., Olsen, S.C., Tumblin, J., Gooch, B.: Color2gray: salience-preserving color removal. ACM Trans. Graph. **24**(3), 634–639 (2005)
14. Grundland, M., Dodgson, N.A.: Decolorize: fast, contrast enhancing, color to grayscale conversion. Pattern Recognit. **40**(11), 2891–2896 (2007)
15. Hanbury, A., Serra, J.: A 3D-polar coordinate colour representation suitable for image analysis. Technical Report PRIP-TR-77 (2002)
16. Kim, Y., Jang, C., Demouth, J., Lee, S.: Robust color-to-gray via nonlinear global mapping. ACM Trans. Graph. **28**(5), 1–4 (2009)
17. Kuhn, G.R., Oliveira, M.M., Fernandes, L.A.F.: An improved contrast enhancing approach for color-to-grayscale mappings. Vis. Comput. **24**, 505–514 (2008)
18. Kuk, J.G., Ahn, J.H., Cho, N.I.: A color to grayscale conversion considering local and global contrast. In: Computer Vision—ACCV 2010, vol. 6495, pp. 513–524 (2011)
19. Lee, H.C., Cok, D.: Detecting boundaries in a vector field. IEEE Trans. Signal Process. **39**(5), 1181–1194 (1991)
20. Lu, C., Xu, L., Jia, J.: Contrast preserving decolorization. In: 2012 IEEE International Conference on Computational Photography (ICCP), pp. 1–7 (2012)
21. Neumann, L., Čadík, M., Nemcsics, A.: An efficient perception-based adaptive color to gray transformation. In: Proceedings of Computational Aesthetics 2007, pp. 73–80 (2007)
22. Pridmore, R.W.: Effects of luminance, wavelength and purity on the color attributes: brief review with new data and perspectives. Color Res. Appl. **32**(3), 208–222 (2007)
23. Prip, A.H., Hanbury, A.: Circular statistics applied to colour images. In: 8th Computer Vision Winter Workshop (2003)
24. Qinghua, H., Pedrycz, W., Yu, D., Jun, L.: Selecting discrete and continuous features based on neighborhood decision error minimization. IEEE Trans. Syst. Man Cybern., Part B, Cybern. **40**(1), 137–150 (2010)
25. Rasche, K., Geist, R., Westall, J.: Re-coloring images for gamuts of lower dimension. Comput. Graph. Forum **24**(3), 423–432 (2005)
26. Setiawan, N., Seok-Ju, H., Jang-Woon, K., Chil-Woo, L.: Gaussian mixture model in improved hls color space for human silhouette extraction. In: Advances in Artificial Reality and Tele-Existence. Lecture Notes in Computer Science, vol. 4282, pp. 732–741. Springer, Berlin (2006)
27. Shapira, L., Shamir, A., Cohen-Or, D.: Image appearance exploration by model-based navigation. Comput. Graph. Forum **28**(2), 629–638 (2009)
28. Shental, N., Bar-hillel, A., Hertz, T., Weinshall, D.: Computing Gaussian mixture models with EM using equivalence constraints. In: Advances in Neural Information Processing Systems, vol. 16. MIT Press, Cambridge (2003)
29. Smith, K., Landes, P.E., Thollot, J., Myszkowski, K.: Apparent greyscale: a simple and fast conversion to perceptually accurate images and video. Comput. Graph. Forum **27**(2), 193–200 (2008)
30. Tanaka, G., Suetake, N., Uchino, E.: Derivation of the analytical solution of color2gray algorithm and its application to fast color removal based on color quantization. Opt. Rev. **16**, 601–612 (2009)
31. Treisman, A., Gelade, G.: A feature-integration theory of attention. Cogn. Psychol. **12**(1), 97–136 (1980)

Author's personal copy

32. Wyszecki, G.: Correlate for lightness in terms of cie chromaticity coordinates and luminous reflectance. J. Opt. Soc. Am. **57**(2), 254–257 (1967)
33. Wyszecki, G.W., Stiles, W.S.: Color Science: Concepts and Methods, Quantitative Data and Formulae. Wiley, New York (1982)
34. Zhao, Y., Tamimi, Z.: Spectral image decolorization. Adv. Vis. Comput. **6454**, 747–756 (2010)
35. Zhu, S.Y., Luo, M.R., Cui, G.H., Li, C.J., Rigg, B.: Comparing large colour-difference data sets. Color Res. Appl. **36**(2), 111–117 (2011)

Wei Zhu is a Ph.D. of the Department of Mathematics, Zhejiang University, China. She received a B.Sc. and Ph.D. in Mathematics from Zhejiang University in 2007 and 2012. Her research interests include image processing and computer graphics.

Ruizhen Hu is a Ph.D. candidate at Zhejiang University in China. She received a B.Sc. in Mathematics from Zhejiang University in 2010. Her research interests are computer graphics and geometry processing.

Ligang Liu is a professor at the School of Mathematical Sciences, University of Science and Technology of China. He received his B.Sc. in applied mathematics (1996) and his Ph.D. in computer graphics (2001) from Zhejiang University, China. Between 2001 and 2004, he worked at Microsoft Research Asia as an associate researcher. Then he worked at Zhejiang University as an associate professor and professor during 2004 and 2012. He paid an academic visit to Harvard University during 2009 and 2011. His research interests include digital geometric processing, computer graphics, and image processing.