Global Color Saliency Preserving Decolorization

Jie Chen\textsuperscript{1}, Xin Li\textsuperscript{1}, Xiuchang Zhu\textsuperscript{1}, Jin Wang\textsuperscript{2}

\textsuperscript{1} Key Lab of Image Processing and Image Communication of Jiangsu Province, Nanjing University of Posts and Telecommunications (NUPT), Nanjing, 210003, China
{hucjjt\texttt{@163.com, lixin@njupt.edu.cn, zhuxc@njupt.edu.cn}  
\textsuperscript{2} College of Information Engineering, Yangzhou University, Yangzhou, China
jinwang@yzu.edu.cn

Abstract. The process of transforming a color image with three channels to a single channel grayscale image is called decolorization, which will unavoidably accompany with the information loss. In this paper, we propose a method to obtain a grayscale image which best preserves the global saliency of the color image. First, we convert the color image to the YUV color space, which separates the luminance and chrominance channels, and conduct parametric linear mapping on Y,U,V channels, through selecting different parameters to get different candidate grayscale images. Then, we compute the global contrast based saliency maps of the color image and candidate grayscale images. Finally, we use the Normalized Cross-Correlation metric to select the grayscale image whose saliency map is the most similar to that of the color image as the wanted decolorization result. The experiment results show that our method retains part of the chrominance information, prevents the contrast degradation in the isoluminance colors, and the global saliency preserving purpose reduces the abrupt change or distortion around the edge areas.

Keywords: Decolorization, Color-to-gray, Global saliency, Parametric linear map, Normalized Cross-Correlation.

1 Introduction

Decolorization is the process of converting a color image to its grayscale form. The most intuitive mode is taking the luminance channel of the color image (e.g. Y channel in the YUV color space) as the decolorization result\textsuperscript{1}. As shown in Fig. 1, this manner cannot keep the contrast of the isoluminance colors. How to keep the structure and contrast of the color image during the color-to-gray conversion is an important research content of decolorization methods.
The purpose of decolorization is to preserve as much visually meaningful information about the color image as possible, meanwhile produce perceptually natural and pleasing grayscale image\cite2. Choosing what features to preserve during the decolorization process is crucial for the achievement of the purpose, and will heavily decide the performance and effect of the method. The work of \cite4 chose the image apparent lightness feature to maintain the perceptual accurate of the conversion. The method of \cite5 defined three visual cues including color spatial consistency, image structure information and color channel perception priority to be the basis of the conversion. \cite6 chose the nature order of hues as the preserved feature of the conversion. The work of \cite7 alleviated the strict color order constraint and aimed at maximally preserving the original color contrast.

Although the main purpose of the decolorization methods is the same, because of the different choice of features in the specific method, there is not a robust method can get satisfied results for all kinds of images, that is each method can be found some failure cases. To solve this problem, the decolorization methods tried to consider multiple features simultaneously which made the algorithms become more and more complex. Nowadays, the research of visual saliency estimation which is relevant to the human visual system has drawn great attention. Saliency originates from visual uniqueness, unpredictability, rarity or surprise, and is often ascribed to variations in image attributes like color, gradient, edges and boundaries\cite3. Using image saliency as the preserved feature in the color-to-gray conversion, can cover multiple attributes of the image simultaneously, and is a comparatively ideal manner in image decolorization. The methods in [1][8-11] all used the image visual saliency as the feature to conduct the color-to-gray conversion. [9] computed the global color saliency of the pixel to instruct the parameter selection in the luminance and chrominance information fusion procedure, this method could enhance the chrominance contrast in the grayscale image, but, it only used few points with the largest saliency value as the instruction, which made it difficult to maintain the contrast of the whole image. The methods proposed in [1][8][10] and [11] computed the region saliency of the image to instruct the parameters selection or adjustment to
convert the image, first step of these methods was image segmentation, which would influence the effect of the methods in the image edge areas.

In this paper, we compute the global saliency value of every pixel without using image segmentation, and use the saliency maps of the color and grayscale images to get the decolorization image. Experimental results show that our method can reduce the contrast loss in the isoluminance colors and prevent the abrupt change or discontinuity around the edge areas.

2 Global color saliency preserving decolorization method

Our decolorization method contains 3 steps, which are parametric linear mapping, global saliency maps computation and similarity measurement. The whole framework is shown in Fig.2.

Fig. 2. Framework of the proposed decolorization method.

2.1 Parametric linear mapping function

Compared with RGB color space, the YUV color space makes the luminance information(Y) separate from the chrominance information(U,V), which is conformable with human visual perception. We convert the input color image to the YUV color space, and operate parametric linear mapping (PLM) on Y,U,V channels.

Because the human visual perception is more sensitive to the contrast of the adjacent pixels than their own values, we alleviate the constraint of the strict color order and define the mapping function as,

\[
G = wY + \alpha_1 U + \alpha_2 (1-U) + \beta_1 V + \beta_2 (1-V) .
\]  

(1)

Where, \(Y, U, V\) respectively is the orderly formed vector of each channel and the values in the vectors are normalized to \([0,1]\), \(w, \alpha_1, \alpha_2, \beta_1, \beta_2\) are the parameters of the
mapping function, with the constraint that $\alpha_i \times \alpha_i = 0$ and $\beta_i \times \beta_i = 0$ to avoid the offset of the two parameterized values in the same chrominance channel. With numbers of different parametric combinations, the mapping process will get numbers of candidate grayscale images, denote the number as $n$. We choose YUV color space to conduct the process, could intuitually supplement the image’s luminance information using its chrominance components.

### 2.2 Computation of the global contrast based visual saliency map

The existing image saliency based or referenced decolorization methods generally use local region saliency of the image, which will be influenced by the region segmentation and edge distortion. In this paper, we take the global contrast based salience estimation (GCSE) algorithm[3] to get the global visual saliency map without conducting image segmentation or edge detection. The original GCSE algorithm is designed to calculate the saliency values of color image, as to our need to calculate for the grayscale image, we make a little adjustment to meet the demand.

The GCSE algorithm computes the color distance $D(c_i, \cdot)$ in the CIELab color space, which is

$$
D(c_i, c_j) = \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} .
$$

(2)

where, $c_i, c_j$ represent two colors in the image, $L_i, a_i, b_i$ respectively represent the three components of color $c_i$ in the CIELab color space.

The number of the colors in the color image is denoted as $N$, the occurring frequency of color $c_i$ is denoted as $f_i$, which can be directly get from the color histogram. Then the saliency value of pixel $P_i$ with color $c_i$ is defined as:

$$
S(P_i) = S(c_i) = \sum_{i=1}^{N} f_i D(c_i, c_i) .
$$

(3)

In consideration of reducing the time complexity, the image colors should go through RGB color space quantization and partial discarding before calculating in equation(3). First, respectively uniform quantizing the R,G,B channel of the image to 12 different values, which will reduce the color number from $256^3$ of the true color space to $12^3=1728$. Then, discarding the low frequency colors, ensuring those reserved high frequency colors could cover more than 95% pixels of the image, and those discarded less than 5% pixels’ colors are replaced by their nearest colors in the reserved color set. Through channel quantizing and color discarding process, the color number $N_c$ generally could be reduced to less than 90.

After substituting the number reduced colors into the equations to get their saliency values, considering the influence of the quantization noise, the algorithm takes a CIELab color space smoothing procedure to refine the saliency values. That is, replacing the color’s saliency value by the weighted average value of its $m$ similar colors’ (measured by CIELab color distance $D(\cdot, \cdot)$) saliency values. Denoting the
color to be smoothed as $c_i$, its similar colors’ number $m$ is setting as a fix value $m = \frac{N}{4}$, the weight value of its similar color $c_j$ is defined as:

$$Weight(c_i, c_j) = T_i - D(c_i, c_j). \quad (4)$$

$$T_i = \sum_{j=1}^{m} D(c_i, c_j). \quad (5)$$

Then, the refined saliency value of color $c_i$ is:

$$S_i(c_i) = \frac{\sum_{j=1}^{m} (T_i - D(c_i, c_j))S(c_j)}{(m-1)T_i}. \quad (6)$$

Where, $(m-1)T_i$ is the normalization factor. After getting saliency values of the image colors, substituting them into the corresponding pixels will get the whole saliency map.

About the calculating of the grayscale image’s saliency map, the procedure is almost the same as the GCSE algorithm, except for changing the color histogram to grayscale histogram, and eliminating the channel quantization process but directly use the 256-level grayscale representation. In view of maintaining the consistency of the color and grayscale saliency maps, the grayscale process still chooses high frequently occurring values which cover more than 95% pixels to replace those less than 5% low frequency values, and after getting the grayscale saliency values, also conducts smoothing to refine the values.

### 2.3 Selecting of the output grayscale image

To get the wanted grayscale image from the $n$ candidates, we take the Normalized Cross-Correlation (NCC) metric to measure the similarity between the color image saliency map and the candidate grayscale images saliency maps.

$$NCC = \frac{\sum_{(x,y)}(S_c(x,y)S_g(x,y))}{\sqrt{\sum_{(x,y)}S_c(x,y)^2\sum_{(x,y)}S_g(x,y)^2}}. \quad (7)$$

Where, $(x,y)$ is coordinate of the pixel in the image. $S_c, S_g$ respectively is the color and grayscale image saliency map. We choose the image which gets the highest NCC value as the output grayscale image.

### 3 Experiment set and analysis

The computation cost of our method mainly lies on the saliency map estimation procedure. In the experiment, we fix factor $w$ to 1, and constraint parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ in the range of $[0,0.5]$ to reduce $n$ and emphasize the importance of luminance among the three channels. And because the empirical result shows that slightly varying the factors would not change the grayscale appearance too much [12],
we discretize the parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ in the range of [0,0.5] with interval 0.1. Considering the constraints of $\alpha_1 \times \alpha_2 = 0$ and $\beta_1 \times \beta_2 = 0$, the number of candidate grayscale images is $n = 121$.

We use Cadik’s decolorization dataset [13] which is the publicly available decolorization benchmark dataset to evaluate our algorithm, the Cadik’s dataset contains 24 different resolution color images. We implement our algorithm in Matlab, for a 390x390 image, it takes about 20 seconds on a computer with a 2.4GHz Intel Core i5 CPU.

We compare our results with CIE Y, Grundland07 [14], Lu12 [7], Liu13 [10] and Du15 [11] results, the images are shown in Fig. 3. It can be seen that our method can get satisfactory results of the 24 images. Especially, our result of image 4 distinguish all the color difference of the balls which is superior to results of CIE Y, Grundland07 [14], Lu12 [7], Liu13 [10], and without excessive enhance the contrast of the two green balls like Du15 [11]. Our results of image 7 and 8 unambiguously differentiate the different color regions which is better than all the other compared methods. Our result of image 22 is the only image obviously preserves the red and green leaves difference on the right side of the color image.

For quantitative evaluation, we employ color contrast preserving ratio (CCPR) metric which was proposed by [7] and widely adopted by the subsequent methods.

$$CCPR = \frac{\# \{ (x, y) \mid (x, y) \in \Omega, |g_x - g_y| \geq r \}}{\# \Omega}. \quad (8)$$
Where, $\Omega$ is the set containing all the pixel pairs $(x, y)$ of the original color image with difference $\delta_{(x,y)} \geq \tau$, $\tau$ is a set threshold, $\|\Omega\|$ is the number of pixel pairs in $\Omega$. $\# \{(x, y) \mid (x, y) \in \Omega, |g_x - g_y| \geq \tau\}$ is the number of pixel pairs in $\Omega$ has difference $|g_x - g_y| \geq \tau$ after decolorization. We evaluate different methods based on CCPR using the 24 color images in Cadik’s dataset, varying $\tau$ from 1 to 40, and get the average CCPRs for the whole dataset as shown in Fig.4.

Although the average CCPR of Du15[11] is higher than our method, but the algorithm is based on the image separation, which will bring abrupt change between different regions, as shown in image 17 of Fig.3, we enlarged the images in Fig.5. And because Du15[11] algorithm takes the middle points of the separate regions to represent every region, it needs a large number of separations to ensure the accuracy of the representation, which will bring high computation cost.

![Fig. 4. CCPRs plot on Cadik’s dataset.](image)

![Fig. 5. Example of Du15[11] algorithm abrupt changes the gradual change colors in the decolorization result, and our algorithm follows the change better.](image)

### 4 Concluding remarks

This paper presents a global saliency based decolorization algorithm to mimic human visual perception and try to preserve the perceive uniformity between the color and grayscale images. We take into account the global saliency values to avoid the influence of local block artifacts and edge distortion. The computation cost of the algorithm has nearly linear correlation with the candidate grayscale images number $n$, how to further reduce $n$ to enhance the algorithm efficiency is a meaningful problem.
Acknowledgments. This work is partially supported by the National Natural Science Foundation of China (61071166, 61172118, 61071091, 61471201), Jiangsu Province Postgraduate Innovative Program of Scientific Research (CXLX12_0474). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References