CorrC2G: Color to Gray Conversion by Correlation

Hossein Ziaei Nafchi, Atena Shahkolaei, Rachid Hedjam, and Mohamed Cheriet, Senior Member, IEEE

Abstract—In this letter, a novel decolorization method is proposed to convert color images into grayscale. The proposed method, called CorrC2G, estimates the three global linear weighting parameters of the color to gray conversion by correlation. These parameters are estimated directly from the correlations between each channel of the RGB image and a contrast image. The proposed method works directly on the RGB channels; it does not use any edge information nor any optimization or training. The objective and subjective experimental results on three available benchmark datasets of color to gray conversion, e.g. Cadik, CSDD and Color250, show that the proposed decolorization method is highly efficient and comparable to recent state-of-the-art decolorization methods. The MATLAB source code of the proposed method is available at: http://www.synchromedia.ca/system/files/CorrC2G.m.

Index Terms—Decolorization, color to gray conversion, correlation, luminance, grayscale, RGB.

I. INTRODUCTION

In many real-world image/video processing and computer vision applications, the 3D color image needs to be transformed into a 1D grayscale image. This is a lossy but a necessary conversion for several applications [1]. Recent years have seen several efforts in developing novel decolorization methods that are more likely to follow human perception of brightness and contrast [2]–[13]. Color to gray (C2G) conversion methods can be categorized into global, local, and hybrid. The global mapping approach has the potential to produce natural looking grayscale outputs. In contrast, local mapping techniques [3], [5] that better preserve the local contrast may produce unnatural outputs. In local mapping methods, the same color pixel within an image might be mapped into different grayscale values, which is generally not desired. Therefore, several methods consider global and local contrast or features for conversion [10], [14], [15]. Besides, video decolorization methods such as [9], [13] are specifically developed in order to maintain temporal coherence of videos.

Since the proposed method belongs to the category of global mapping, we focus on these methods. Gooch et al. [2] proposed a method to maintain color contrast between pixel pairs by optimizing an objective contrast function. Kim et al. [9] proposed a non-linear parametric model in which the parameters are estimated by minimizing an objective function that preserves color differences. In several recent global mapping methods, the input color image I is converted into a grayscale output g by linear weighting of the R, G, and B channels, i.e. \( g(i, j) = \sum_{c=R,G,B} \lambda_c I_c(i, j) \), where \( \sum_{c=R,G,B} \lambda_c = 1 \). Here, the three linear weighting parameters \( \lambda \) should be estimated on the basis of some models. In [7], a gradient error energy function is minimized to compute the three linear weighting parameters. This interesting approach was given notable consideration and some variations of this method has been proposed [11], [12]. While the method of [11] objectively preserves the contrast and run in real-time, it may produce grayscale outputs with an unnatural appearance. In contrast, the method proposed in [12] produces mostly natural outputs but at the cost of being several times slower.

In this letter, we propose a novel decolorization method that estimates the three global linear weighting parameters \( \lambda \) directly from the R, G, and B channels. The correlations between each channel of the color image with a base image map, which is very likely to preserve contrast, are mapped to \( \lambda \). To the best of our knowledge, correlation has not been used for the purpose of C2G conversion. The proposed method takes into account both the magnitude and sign of the correlation values to adjust the weighting parameters. The proposed training-free method is very simple, it runs in real-time and offers perceptually consistent grayscale outputs with good contrast preservation.

II. PROPOSED DECOLORIZATION METHOD

The proposed decolorization method is a global mapping approach that estimates the three linear weighting parameters \( \lambda \) from correlation. Correlation is a measure of association between variables [16]. Here, we use Pearson’s measure of correlation between two variables X and Y, which is commonly defined as:

\[
\rho_{X,Y} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2}(\sum(Y_i - \bar{Y})^2)^{1/2}}
\]  

where, \( \bar{X} \) and \( \bar{Y} \) are means of variables X and Y, respectively. It is worth noting that an equivalent formula for \( \rho \) is \( s_{XY}/s_Xs_Y \), where \( s_{XY} \) is the sample covariance, and \( s_X \) and \( s_Y \) are sample standard deviations. Given the R, G, and B channels of a color image, the correlation between each channel with an image map that is likely to preserve contrast is computed and normalized in order to estimate the three weighting parameters \( \lambda \).

In order to produce a 2D image which reflects the contrast of the color image, the following two images are combined. The first image is simply the mean image \( \mu \), and the second is the standard deviation image \( \sigma \):

\[
\mu_g(\mathbf{x}) = \frac{1}{3} (R_g(\mathbf{x}) + G_g(\mathbf{x}) + B_g(\mathbf{x}))
\]

\[
\sigma_g(\mathbf{x}) = \sqrt{(R_g(\mathbf{x}) - \mu_g(\mathbf{x}))^2 + (G_g(\mathbf{x}) - \mu_g(\mathbf{x}))^2 + (B_g(\mathbf{x}) - \mu_g(\mathbf{x}))^2}
\]

The proposed decolorization method directly estimates the weighting parameters from the correlations between each channel of the color image and a map that is likely to preserve contrast. The weighting parameters are scaled by the mean and the standard deviation of the grayscale output to produce a grayscale image with good contrast.

This method is expected to produce good quality grayscale outputs with good contrast, whereas the method of [11], which uses the gradient energy function, produces mostly natural outputs but at the cost of being several times slower.
\[ \mu(i, j) = \frac{1}{3} \sum_{c=R,G,B} I_c(i, j) \] (2)

\[ \sigma(i, j) = \left[ \frac{1}{2} \sum_{c=R,G,B} |I_c(i, j) - \mu(i, j)|^{\frac{2}{\gamma}} \right]^{1/2} \] (3)

The values of \( \sigma \) are further divided by 147,2243, which is the maximum possible value of \( \sigma \). In what follows, the rationale to use the two images, \( \mu \) and \( \sigma \), and their combination are explained. Considering the \([0 255]\) range for the 3D RGB image, there are \(256^3 = 16,777,216\) possible color values. However, possible values for the 2D images \( \mu \) and \( \sigma \) (1D vectors) are just 766 and 16,365, respectively. This clearly indicates the probability of contrast loss as a result of the 3D to 1D conversion. In order to reduce the probability of contrast loss, we propose to use the pointwise product of \( \mu \) and \( \sigma \), which provides more than 2 million possible values:

\[ Q(i, j) = \mu(i, j) \times \sigma(i, j) \] (4)

Fig. 1 shows a color image with its mean \( \mu \), standard deviation \( \sigma \), and contrast map \( Q \).

The three Pearson correlation values between RGB channels and \( Q \) are denoted as: \( P = \{ \rho_{RQ}, \rho_{GQ}, \rho_{BQ} \} \). Each correlation value lies in \([-1 1]\). Given these three correlations, the purpose is to map them to the three weighting parameters, i.e. \( P \rightarrow \lambda \). For mapping, the unsigned correlations (absolute values) are mapped to \( \beta \), and original values of the correlations are mapped to \( \gamma \). Finally, \( \lambda \) is computed from \( \beta \) and \( \gamma \). The reason for the consideration of the absolute correlations along with the signed correlations is to avoid mapping larger negative correlations to the smaller weighting parameters. The parameters of \( \lambda \) are computed using the following simple calculations. The first assumption is that a channel with a higher correlation with the contrast map \( Q \) should take a larger weighting parameter:

\[ \beta_c = \frac{|P_c|}{\sum |P|} \] (5)

where, \( c \) is the channel index and \( \sum \beta = 1 \). At the same time, a channel with inverse correlation with \( Q \) should take a lower weighting parameter:

\[ \gamma_c = \frac{P_c - \min_P}{\max_P - \min_P} - 0.5 \] (6)

where, \( \gamma \in [-0.5, 0.5] \). In the above equation, the constant 0.5 can be replaced with any other value in range \([0, 1]\) to control the contribution of the inverse correlations. \( \lambda \) is estimated by adding \( \min(\beta, \gamma) \) to \( \beta \):

\[ \lambda_c = |\beta_c + \min(\beta_c, \gamma_c)| \] (7)

Finally, \( \lambda \) is normalized so that \( \sum_{c=R,G,B} \lambda_c = 1 \). In the above equation, \( \gamma \) is used only when it decreases weighting parameters. We recall that \( \gamma \) was used to decrease the weighting parameters of those channels with inverse correlations. In contrast to the other methods that estimate the weighting parameters \( \lambda \) by optimizing an objective function \([2, 7, 11, 12]\), the proposed method directly estimates \( \lambda \) from the correlation values. The problem with such methods is that the defined objective function does not necessarily follow human perception of brightness and contrast.

![Fig. 2. Two possible grayscale outputs of the proposed method for a color image. (b) is produced by using the standard deviation image \( \sigma \), and (c) is produced by using the complement of \( \sigma \).](image)

Depending on the preference of users for the perceived color, more saturated colors are perceived to be either brighter or darker than their luminance \([6]\). In this regard, the standard deviation image \( \sigma \) can be replaced with its complement image \((1 - \sigma)\) and parameters of \( \lambda \) can be estimated accordingly. In this approach, the proposed method has two grayscale outputs. Fig. 2 gives an example of these two grayscale outputs. Some users may prefer one or the other of them, while others may evaluate them as equal. Since the proposed method should produce a single output, the one with more and larger peaks at the middle of its histogram is chosen as the final output. This two-output strategy slightly improves the objective and subjective results. The objective performance can be greatly improved if the final output of the method is chosen according to the C2G evaluation metrics, such as the E-score \([17]\) and C2G-SSIM \([18]\). We did not use these metrics because they are several times slower than the histogram analysis approach.

### III. Experimental Results

In the experiments, three available datasets for evaluation of the color to gray methods are used. The Cadik dataset \([19]\) contains 25 (originally 24) saturated color images. The Color250 dataset \([17]\) comprises 250 color images with a wide range of natural and synthetic images. The third dataset is CSDD \([10]\), which contains 22 color images with abundant colors and patterns. For objective evaluation, two objective quality assessment metrics for color to gray image conversion are used: E-score \([17]\) and C2G-SSIM \([18]\). E-score evaluates both the color contrast preservation ratio (CCPR) and color content fidelity ratio (CCFR). C2G-SSIM is a more recent C2G evaluation metric based on the popular image quality assessment metric SSIM \([20]\). In comparison with the E-score, C2G-SSIM provides a useful quality map and shows higher correlation with human subjective evaluations. Five state-of-the-art C2G methods were chosen for comparison \([4, 7, 10–12]\). Each of these methods has shown very promising...
In Fig. 3, outputs of the six C2G methods for eight color images are shown. Here, $r$ is a downsampling parameter (see subsection Complexity). It can be seen from Fig. 3 that the proposed method fairly shows the color differences.

According to the results of Table I, GcsDecolor provides highest performance based on E-score, and the proposed method shows highest performance based on C2G-SSIM. In terms of E-score, the proposed method shows better results than the Decolorize, Saliency, and SPDecolor methods.

It is common to report qualitative performance based on the CCPR by varying its parameter $\tau$, a threshold below which the color differences become almost invisible to the human visual system [17]. Fig. 4 shows the results for six C2G methods. We can see that GcsDecolor yields the best results, and that the proposed method is comparable with the other methods.

We also conducted a subjective evaluation in which the outputs of the proposed C2G method are compared with those of other methods. In this experiment, three subjects with a background in image processing were asked to discuss why they preferred the output of one method to that of another. After discussion, only one vote was given for each pair-of-method comparison.
CorrC2G (r = 512) for 127 images, they are rated equal for 139 images with a factor of $\tau = 256$. We can see that the proposed method shows better performance than the other methods. For example, outputs of GcsDecolor are rated worse than the proposed method CorrC2G (r = 256) for 127 images, they are rated equal for 139 images, and better for 31 images. GcsDecolor, which has the highest E-score performance, shows the worst results on the basis of subjective evaluation. In fact, the correlation between C2G-SSIM and the subjective evaluations is higher than that for the E-score.

**TABLE II**

<table>
<thead>
<tr>
<th>C2G method</th>
<th>CorrC2G (r = 512)</th>
<th>CorrC2G (r = 256)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>worse</td>
<td>equal</td>
</tr>
<tr>
<td>Decolorize</td>
<td>81</td>
<td>166</td>
</tr>
<tr>
<td>RTCP</td>
<td>69</td>
<td>200</td>
</tr>
<tr>
<td>Saliency</td>
<td>91</td>
<td>168</td>
</tr>
<tr>
<td>GcsDecolor</td>
<td>129</td>
<td>140</td>
</tr>
<tr>
<td>SPDecolor</td>
<td>49</td>
<td>209</td>
</tr>
</tbody>
</table>

**TABLE III**

<table>
<thead>
<tr>
<th>C2G method</th>
<th>128 $\times$ 128</th>
<th>384 $\times$ 512</th>
<th>1080 $\times$ 1920</th>
<th>2160 $\times$ 3840</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decolorize</td>
<td>4.28</td>
<td>65.83</td>
<td>771.15</td>
<td>3078.83</td>
</tr>
<tr>
<td>RTCP</td>
<td>11.87</td>
<td>19.04</td>
<td>67.20</td>
<td>219.96</td>
</tr>
<tr>
<td>Saliency</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GcsDecolor</td>
<td>16.43</td>
<td>25.40</td>
<td>73.28</td>
<td>226.97</td>
</tr>
<tr>
<td>SPDecolor</td>
<td>25.51</td>
<td>225.97</td>
<td>7883.76</td>
<td></td>
</tr>
<tr>
<td>CorrC2G (r = 512)</td>
<td>2.87</td>
<td>37.67</td>
<td>138.18</td>
<td>230.36</td>
</tr>
<tr>
<td>CorrC2G (r = 256)</td>
<td>2.87</td>
<td>16.76</td>
<td>50.46</td>
<td>147.46</td>
</tr>
<tr>
<td>CorrC2G (r = 64)</td>
<td>2.78</td>
<td>5.60</td>
<td>35.16</td>
<td>134.37</td>
</tr>
</tbody>
</table>

**IV. Conclusion**

This letter introduces a novel correlation-based decolorization method to convert color images into grayscale. The Pearson correlations between channels of color images with a contrast map are mapped directly to the three linear weighting parameters. The proposed method is very simple and runs in real-time, yet it offers perceptually consistent outputs with a natural appearance. Extensive objective and subjective experiments on the three benchmark datasets verified the efficiency of the proposed method. Several possible ways of improving the current method would be to use a better numerically compatible contrast map, more accurate mapping of the correlations to the three weighting parameters, and to modify the double-output behavior of the proposed method.

**ACKNOWLEDGMENTS**

The authors thank the NSERC of Canada for their financial support under Grants RGPDD 451272-13 and RGPIN 138344-14.
REFERENCES


