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. [6] proposed a non-linear parametric model in which the parameters are estimated by minimizing an objective

H. Ziaei Nafchi, A. Shahkolaei and M. Cheriet are with the Synchromedia Laboratory for Multimedia Communication in Telepresence, École de technologie supérieure, Montreal (QC), Canada H3C 1K3 (email: hossein.zi@synchromedia.ca; atena.shahkolaei.1@ens.etsmtl.ca; mohamed.cheriet@etsmtl.ca

R. Hedjam is with the Department of Geography, McGill University, 805 Sherbrooke Street West, Montreal, QC H3A 2K6, Canada (email: rachid.hedjam@mcgill.ca)

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 λ, 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 λ 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 λ. 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 λ 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:

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

where, $\bar{X}$ and $\bar{Y}$ are means of variables X and Y, respectively. It is worth noting that an equivalent formula for ρ is $s_{XY}/s_X s_Y$, 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 λ.

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 μ, and the second is the standard deviation image σ:

$$\mu(i,j) = \frac{1}{3} \sum_{c=R,G,B} \mathrm{I}_c(i,j) \qquad (2)$$

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

The values of σ are further divided by 147.2243, which is the maximum possible value of σ. In what follows, the rationale to use the two images, μ and σ, 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 μ and σ (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 μ and σ, which provides more than 2 million possible values:

$$\mathrm{Q}(i,j) = \mu(i,j) \times \sigma(i,j) \qquad (4)$$

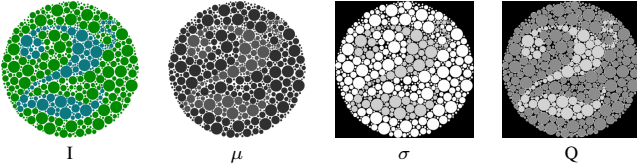Fig. 1 shows a color image with its mean μ, standard deviation σ, and contrast map Q.

Fig. 1. An example of the mean image μ, standard deviation image σ, and contrast map Q for a color image I.

The three Pearson correlation values between RGB channels and Q are denoted as: $\mathcal{P} = \{\rho_{RQ}, \rho_{GQ}, \rho_{BQ}\}$. Each correlation value lies in [-1 1]. Given these three correlations, the purpose is to map them to the three weighting parameters, i.e. $\mathcal{P} \longrightarrow \lambda$. For mapping, the unsigned correlations (absolute values) are mapped to β, and original values of the correlations are mapped to γ. Finally, λ is computed from β and γ. 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 λ 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:

$$\beta_c = \frac{|\mathcal{P}_c|}{\sum |\mathcal{P}|} \qquad (5)$$

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:

$$\gamma_c = \frac{\mathcal{P}_c - min_{\mathcal{P}}}{max_{\mathcal{P}} - min_{\mathcal{P}}} - 0.5 \qquad (6)$$

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. λ is estimated by adding $min(\beta, \gamma)$ to β:

$$\lambda_c = |\beta_c + min(\beta_c, \gamma_c)| \qquad (7)$$

Finally, λ is normalized so that $\sum_{c=R,G,B} \lambda_c = 1$. In the above equation, γ is used only when it decreases weighting parameters. We recall that γ was used to decrease the weighting parameters of those channels with inverse correlations. In contrast to the other methods that estimate the weighting parameters λ by optimizing an objective function [2], [7], [11], [12], the proposed method directly estimates λ 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.
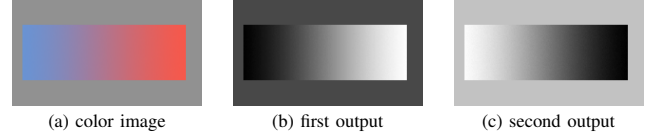

(a) color image   (b) first output   (c) second output

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

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 σ can be replaced with its complement image $(1 - \sigma)$ and parameters of λ 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
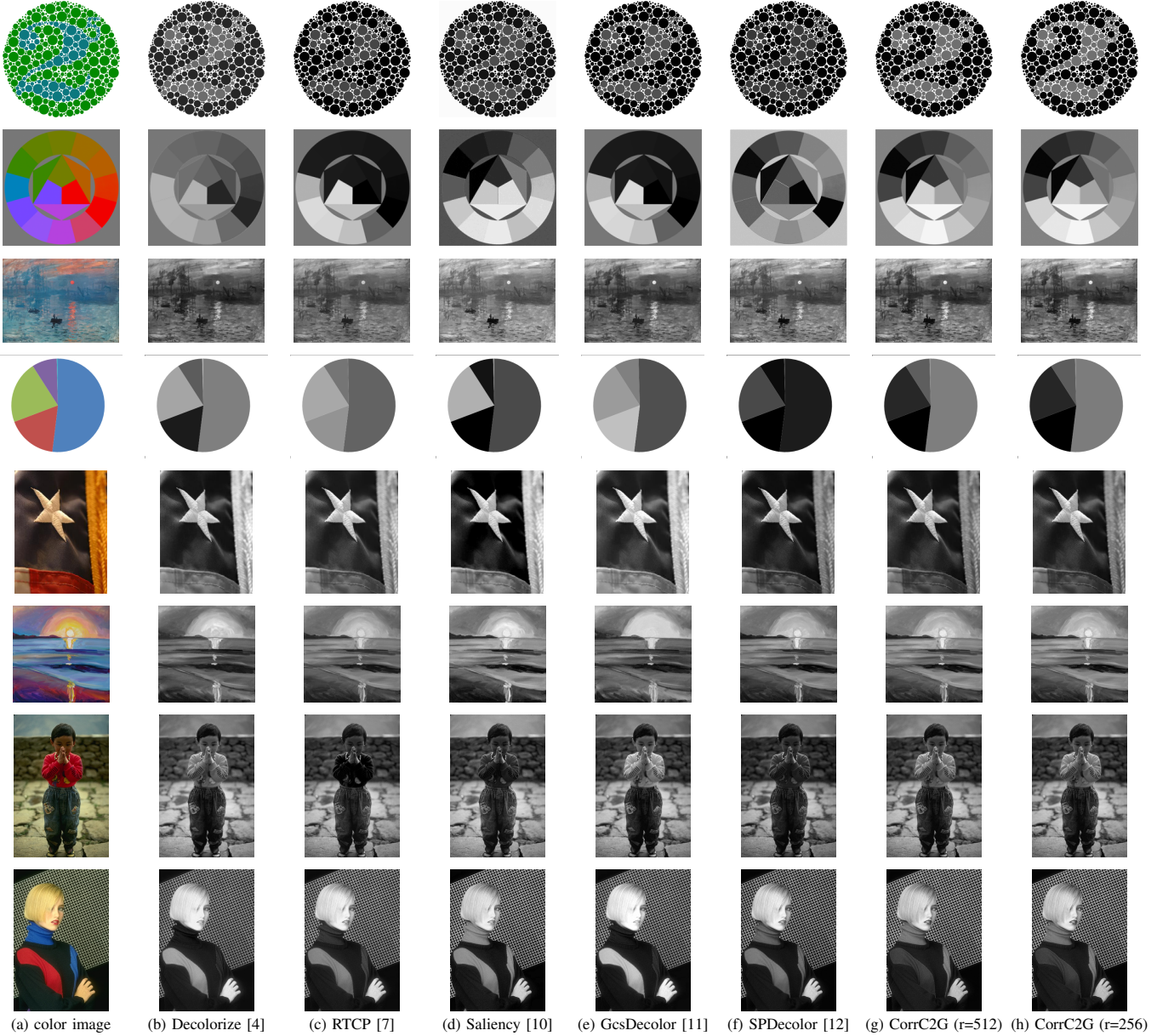
Fig. 3. Visual comparison of six color to gray conversion methods. To view finer detail, please zoom in on the electronic version.

performance in comparison to the other existing methods. 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 τ, 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.

TABLE I
THE AVERAGE PERFORMANCE OF SIX C2G METHODS FOR 297 IMAGES

C2G method (297 images)	E-score ($\tau = 15$)	C2G-SSIM
Decolorize [4]	0.8972	0.8639
RTCP [7]	0.9115	0.8770
Saliency [10]	0.8965	0.8705
GcsDecolor [11]	**0.9162**	0.8707
SPDecolor [12]	0.8952	0.8775
CorrC2G ($r = 512$)	0.8981	**0.8796**
CorrC2G (default, $r = 256$)	0.8987	**0.8796**
CorrC2G ($r = 128$)	0.8957	0.8774
CorrC2G ($r = 64$)	0.8944	0.8777

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-
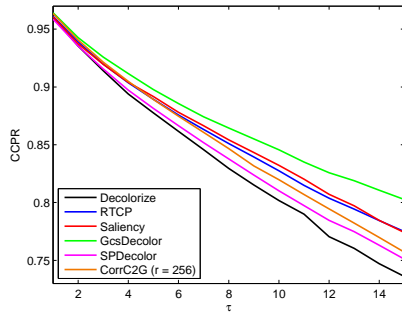
Fig. 4. Comparison of six C2G methods based on the CCPR metric.

not affect the visual appearance of the outputs. Fig. 5 verifies this behavior for eight images.

TABLE III
RUN TIME COMPARISON OF C2G METHODS IN TERMS OF MILLISECONDS

C2G method	128×128	384×512	1080×1920	2160×3840
Decolorize [4]	4.28	65.63	771.15	3078.83
RTCP [7]	11.87	19.04	67.20	219.96
Saliency [10]	-	-	-	-
GcsDecolor [11]	16.43	25.40	73.28	226.97
SPDecolor [12]	25.51	225.97	2239.70	7883.76
CorrC2G ($r = 512$)	2.87	37.67	138.18	230.36
CorrC2G ($r = 256$)	2.87	16.76	50.46	147.46
CorrC2G ($r = 128$)	2.87	7.45	39.75	137.41
CorrC2G ($r = 64$)	2.78	5.60	35.16	134.37

comparison: 'worse', 'equal', or 'better'. The subjects had no prior knowledge of the compared methods and that the grayscale pairs were randomly placed on the screen. In this experiment, 297 color images of the three datasets were considered. In total, 2970 comparisons were recorded. The results of the five C2G methods against the proposed method are listed in Table II. 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
RESULTS OF SUBJECTIVE EVALUATION FOR FIVE C2G METHODS AGAINST THE PROPOSED METHOD FOR 297 COLOR IMAGES

C2G method	CorrC2G (r = 512)			CorrC2G (r = 256)		
	worse	equal	better	worse	equal	better
Decolorize [4]	81	166	50	80	162	55
RTCP [7]	69	200	28	68	198	31
Saliency [10]	91	168	38	91	167	39
GcsDecolor [11]	129	140	28	127	139	31
SPDecolor [12]	49	209	39	50	206	41

A. Complexity

To show the efficiency of the proposed method, a run-time comparison between six C2G methods was performed and is shown in Table III. The experiments were performed on a Corei7 3.40 GHz CPU with 16 GB of RAM. The proposed method was implemented in MATLAB 2013b running on Windows 7. The proposed method first downsamples the color image with a factor of $f = r/min(h, w)$, where h and w are image height and width, and constant $r = 256$ is used by default. Then, it estimates the weighting parameters from the downsampled image. It can be seen from Table III that CorrC2G runs faster than the other methods for images with different resolution. Also, the proposed method runs faster by reducing its downsampling parameter r. This speedup is smaller for larger images because the majority of the run-time is spent on the common operations that are dependent to the image size but independent from the value of r. For majority of the images, reducing the value of r to some threshold does

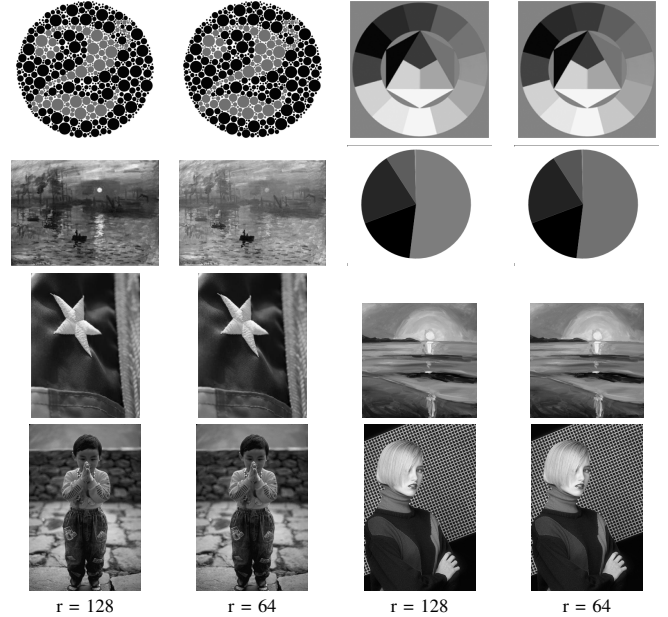

Fig. 5. Outputs of the proposed method given the color images of Fig. 3 for r = 128 (first and third columns) and r = 64 (second and fourth columns). Except for image 'sunrise', other outputs are quite similar.

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.

REFERENCES

[1] C. Kanan and G. W. Cottrell, "Color-to-grayscale: Does the method matter in image recognition?" *PLOS ONE*, vol. 7, no. 1, pp. 1–7, 01 2012.

[2] A. A. Gooch, S. C. Olsen, J. Tumblin, and B. Gooch, "Color2gray: Salience-preserving color removal," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 634–639, Jul. 2005.

[3] L. Neumann, M. Čadík, and A. Nemcsics, "An efficient perception-based adaptive color to gray transformation," in *Proceedings of the Third Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*, ser. Computational Aesthetics'07, 2007, pp. 73–80.

[4] M. Grundland and N. A. Dodgson, "Decolorize: Fast, contrast enhancing, color to grayscale conversion," *Pattern Recognition*, vol. 40, no. 11, pp. 2891 – 2896, 2007.

[5] K. Smith, P. Landes, J. Thollot, and K. Myszkowski, "Apparent greyscale: a simple and fast conversion to perceptually accurate images and video," *Comput Graph Forum*, vol. 27, no. 2, pp. 193–200, 2008.

[6] Y. Kim, C. Jang, J. Demouth, and S. Lee, "Robust color-to-gray via nonlinear global mapping," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 161:1–161:4, Dec. 2009.

[7] C. Lu, L. Xu, and J. Jia, "Real-time contrast preserving decolorization," in *SIGGRAPH Asia 2012 Technical Briefs*, ser. SA '12. New York, NY, USA: ACM, 2012, pp. 34:1–34:4.

[8] Y. Song, L. Bao, X. Xu, and Q. Yang, "Decolorization: Is rgb2gray() out?" in *SIGGRAPH Asia 2013 Technical Briefs*, ser. SA '13. New York, NY, USA: ACM, 2013, pp. 15:1–15:4.

[9] Y. Song, L. Bao, and Q. Yang, "Real-time video decolorization using bilateral filtering," in *IEEE Winter Conference on Applications of Computer Vision*, March 2014, pp. 159–166.

[10] H. Du, S. He, B. Sheng, L. Ma, and R. W. H. Lau, "Saliency-guided color-to-gray conversion using region-based optimization," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 434–443, Jan 2015.

[11] Q. Liu, P. X. Liu, W. Xie, Y. Wang, and D. Liang, "Gcsdecolor: Gradient correlation similarity for efficient contrast preserving decolorization," *IEEE Transactions on Image Processing*, vol. 24, no. 9, pp. 2889–2904, Sept 2015.

[12] Q. Liu, P. Liu, Y. Wang, and H. Leung, "Semi-parametric decolorization with laplacian-based perceptual quality metric," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, 2016.

[13] Y. Tao, Y. Shen, B. Sheng, P. Li, and R. W. H. Lau, "Video decolorization using visual proximity coherence optimization," *IEEE Transactions on Cybernetics*, vol. PP, no. 99, pp. 1–14, 2017.

[14] J. G. Kuk, J. H. Ahn, and N. I. Cho, *A Color to Grayscale Conversion Considering Local and Global Contrast*. Asian Conference on Computer Vision, 2011, pp. 513–524.

[15] Z. Jin, F. Li, and M. K. Ng, "A variational approach for image decolorization by variance maximization," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 944–968, 2014.

[16] J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician*, vol. 42, no. 1, pp. 59–66, 1988.

[17] C. Lu, L. Xu, and J. Jia, "Contrast preserving decolorization with perception-based quality metrics," *International Journal of Computer Vision*, vol. 110, no. 2, pp. 222–239, 2014.

[18] K. Ma, T. Zhao, K. Zeng, and Z. Wang, "Objective quality assessment for color-to-gray image conversion," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4673–4685, Dec 2015.

[19] M. Čadík, "Perceptual evaluation of color-to-grayscale image conversions," *Computer Graphics Forum*, vol. 27, no. 7, pp. 1745–1754, 2008.

[20] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.