doi:10.3772/j.issn.1006-6748.2014.03.006

Colorization method based on the linear relationship assumption①

Nie Dongdong (聂栋栋) *, Wang Li *, Ma Qinyong② **
(* College of Sciences, Yanshan University, Qinhuangdao 066004, P. R. China)
(** College of information Science and Engineering, Yanshan University, Qinhuangdao 066004, P. R. China)

Abstract

A new local cost function is proposed in this paper based on the linear relationship assumption between the values of the color components and the intensity component in each local image window, then a new quadratic objective function is derived from it and the globally optimal chrominance values can be computed by solving a sparse linear system of equations. Through the colorization experiments on various test images, it is confirmed that the colorized images obtained by our proposed method have more vivid colors and sharper boundaries than those obtained by the traditional method. The peak signal to noise ratio (PSNR) of the colorized images and the average estimation error of the chrominance values relative to the original images also show that our proposed method gives more precise estimation than the traditional method.

Key words: colorization, quadratic objective function, optimization, least square solver, color propagation

0 Introduction

Color provides powerful information for image analysis and pattern recognition. Colorizing grayscale images is a useful technique in scientific computing visualization and entertainment. Coloring of photographs by hand is nearly as old as photography itself. The information content of some scientific images can be perceptually enhanced with color by exploiting variations in chromaticity as well as luminance.

Colorization, a computer-assisted image processing method, was firstly introduced by Wilson Markle in 1970 describing the computer-assisted process he invented to add color to black and white movies or TV programs. The term is now generically used to describe any technique for adding color to grayscale images or videos. Traditional colorization is usually carried out manually or assisted by some drawing software tools, which is very tedious and time-consuming, thus automatic or semi-automatic colorization methods are proposed by various researchers in recent years.

To overcome the ambiguity in colorization process, some color hints provided by the user via scribbles or seed pixels are used as initial color conditions and various assumptions are also introduced in the study of different researchers. We assume that there is a linear relationship between the values of the color components and the intensity component in each local window of the images; similar assumptions are widely used in a variety of image analysis problems, such as He Kaiming[1] uses it to refine the transmission map of outdoor haze images and Anat Levin[2] uses it to compute high quality mattes for natural images. Based on the assumption, we have succeeded in turning the under-constrained colorization problem to a minimizing problem of a new cost function, and by eliminating the linear parameters used in each local window, we have obtained a quadratic objective function only in color component values, which can be solved using traditional least square method easily.

The contribution of our method is that we exploit a new local cost function to eliminate linear parameters and propose a new global objective function to estimate the optimal U, V component values under smoothness constraint and a relaxed initial color constraint, which improves the images colorized, and make the image region boundaries sharper with less unsuitable color blur, especially in the regions with complex structures.

1 Previous work

The colorization method has been studied extensively in recent years. We only focus on the work directly related to the work here to describe.

① Supported by the National Natural Science Foundation of China (No. 61073089), the Joint Funds of the National Natural Science Foundation of China (No. U1304616) and the Qinhuangdao Research & Development Program of Science & Technology (No. 2012021A044).

② To whom correspondence should be addressed. E-mail: mqyray@163.com

Received on Mar. 21, 2013

Since 1970's, colorization has been extensively studied in the movie industry. Various semi-automatic analogue techniques[3] have been used to accomplish this challenging task. Gonzalez[4] uses luminance keying to transfer color into the grayscale image. Their approach exploits the user defined look-up table in which each level of grayscale intensity is corresponding to the specified hue, saturation and brightness. This method is extensively used in some commercial post-production systems to apply color on the grayscale image. However, luminance keying cannot resolve the problem when one wants to apply different colors at the same intensity level.

Shi[5] colorizes the grayscale image by segmentation and color filling method, where the image is first segmented into regions and then the desired colors are used to fill each region. Since the existing image segmentation algorithms usually cannot segment the image into meaningful regions, just filling color in each segmented region cannot output natural colorized results. Sykora[6] suggests using unsupervised image segmentation in cartoons colorization. Cartoon generally includes two layers: background and foreground. The foreground layer contains homogeneous regions surrounded by outlines. The background layer remains static during the animation. Taking the advantage of this property, the original grayscale image can be divided into a few regions by a robust outline detector. However the method usually cannot get ideal result for other type of images.

Welsh[7] proposes a colorization method by using local luminance distribution as textural information and transfers the color between similar texture regions in the source color image and the target grayscale image. In their method, texture matching, which has the same properties as luminance keying, replaces the single intensity level comparing. Chen[8] combines grayscale image matting algorithm with color transferring techniques. Since image matting algorithm is effective to handle objects with intricate and vision sensitive boundaries, their method gives exciting results especially for some difficult cases, such as human faces or images with confusing luminance distribution. Lipowezky[9] introduces Bayesian texture classification into their colorization of grayscale aerial or space imagery. By using a prototype matching their method overcome some disadvantages in the method of Welsh et al. The method is also used for BW image transform into the images of different spectra such as color infrared (CIR) image. Charpiat[10] proposes a global optimization method for automatic color assignment. The success of these methods depends heavily on finding a suitable reference image. To avoid this problem, Chia[11] proposes a colorization system to colorize images by using internet images. The user only needs to provide a semantic text label and segmentation cues for major foreground objects in the scene.

Some colorization methods rely on color information provided by the user via scribbles or seed pixels. Horiuchi[12,13] differentitiats among adjacent four-connected pixels. Levin[14] solves colorization as an optimization problem by a few color scribbles user provided. Noda[15] formulates the colorization problem as the maximum a posteriori (MAP) estimation of a color image given a monochrome image. Sapiro[16] obtains image colors by solving a partial differential equation in which the color gradient information is constrained by the monochrome image geometry. Yatziv[17] solves the colorization problem based on the concept of color blending derived from a weighted distance function. Their algorithm is robust to different sets of scribbles. Lagodzinski[18] presents a colorization method based on an hybrid morphological distance transformation which considers the image structures to automatically propagate the color.

Based on the previous work, we develop our image colorization method mainly from the following two aspects:

(1) Define a new cost function based on the local linear relationship assumption between the values of the color components and the intensity component in each local image window, by which our colorization method restricts the relationship between all pixel pairs.

(2) Succeed in turning an ill-posed problem to an optimization problem of a new quadratic objective function composed of a smoothness term and a data term constrains.

2 Our colorization method

The problem of colorization is how to estimate the U and V component values of a grayscale image from the only known intensity value Y with some initial color conditions provided by the user via scribbles or seed pixels.

Because the method to obtain the V component values is similar to get the U component values, we only explain our colorization method using U component values as an instance.

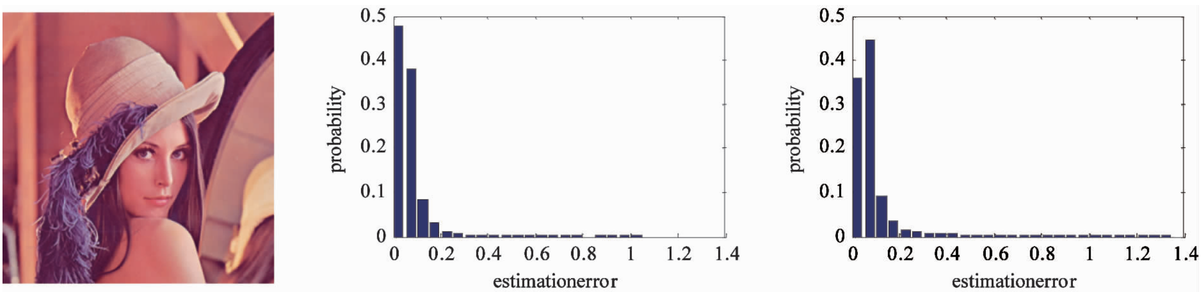
Since colorization is an ill-posed problem, many assumptions are introduced in the previous work to overcome the ambiguity in colorization process. For example, Welsh[7] assumes similar texture regions between the source color image and the target grayscale image have similar colors, and Levin[14] premises that nearby pixels in space that have similar grey levels also have similar colors. We assume that there is a linear relationship between the values of the color components and the intensity component in each local image window.

Taking the most commonly used standard image 'Lena' (see Fig. 1(a)) as an example, it is estimated that the U, V values in each non-overlapped 3×3 windows are based on the linear relationship assumption, and then the relative estimation error between the original and the estimated values of every pixels is computed. The probability distributions of the average relative estimation errors of U, V in each window are shown in Fig. 1(b), (c) respectively. It can be see that in most cases the relative estimation errors are small than 0.1, which confirms our linear relationship assumption is reasonable.

Based on the assumption, in a small window around pixel k, the U component values of each pixel can be expressed by

$$U_i = a_k Y_i + b_k, \quad i \in w_k \tag{1}$$

where w_k is a small window around pixel k, which is usually set to be 3×3 in our experiments; U_i, Y_i is the values of the U component and the Y component of pixel i; a_k, b_k are the parameters to describe the linear relationship of U_i, Y_i in the window w_k.



(a) 'Lena' image (b) Average relative estimation errors of U (c) Average relative estimation errors of V

Fig. 1 Probability distributions of the average relative estimation errors of U, V in each non-overlapped 3×3 windows

In order to eliminate the unknown parameters a_k, b_k used in each window w_k, the following local cost function is introduced:

$$C_k(a_k, b_k) = \sum_{i \in w_k} ((U_i - (a_k Y_i + b_k))^2 + \varepsilon a_k^2) \tag{2}$$

The cost function above includes a regularization term εa_k^2 for getting a smoother solution of U component, since $a_k = 0$ means that the U component values in the window w_k are constant.

Define two matrixes:

$$\boldsymbol{F}_k = \begin{bmatrix} Y_{k_1} & 1 \\ Y_{k_2} & 1 \\ \vdots & \vdots \\ Y_{k_n} & 1 \\ \sqrt{\varepsilon} & 0 \end{bmatrix}, \quad \boldsymbol{G}_j = \begin{bmatrix} U_{k_1} \\ U_{k_2} \\ \vdots \\ U_{k_n} \\ 0 \end{bmatrix}$$

where n is the number of pixels in the window w_k.

Then, the optimal parameters a_k^*, b_k^* inside each window w_k is the solution of a least squares problem:

$$\begin{aligned}(a_k^*, b_k^*) &= \operatorname{argmin} \left\| \boldsymbol{F}_k \begin{bmatrix} a_k \\ b_k \end{bmatrix} - \boldsymbol{G}_k \right\|^2 \\ &= (\boldsymbol{F}_k^T \boldsymbol{F}_k)^{-1} \boldsymbol{F}_k^T \boldsymbol{G}_k \end{aligned} \tag{3}$$

Define matrix:

$$\boldsymbol{H}_k = I - \boldsymbol{F}_k (\boldsymbol{F}_k^T \boldsymbol{F}_k)^{-1} \boldsymbol{F}_k^T \tag{4}$$

where I represents a $(n+1) \times (n+1)$ identity matrix.

Based on Eqs(2), (3), (4), the local quadratic cost function in the U component values can be rewritten as

$$C_k = \left\| F_k \begin{bmatrix} a_k^* \\ b_k^* \end{bmatrix} - G_k \right\|^2 = G_k^T H_k^T \boldsymbol{H}_k \boldsymbol{G}_k \tag{5}$$

The aim of our colorization method is to minimize the costs for all of the local windows under the initial data constraints. Therefore, we adopt the following global objective function to estimate the U component values of a grayscale image:

$$J(U) = \sum_k C_k + \lambda (DU - U_0)^{\mathrm{T}} (DU - U_0) \tag{6}$$

Suppose the image size is $M \times N$, U is a $MN \times 1$ vector representing the U component values of all image pixels to be computed; D is a $MN \times MN$ diagonal matrix and each diagonal element d_{ii} is either 1 when the U component value of pixel j is known or 0 otherwise; U_0 is the $MN \times 1$ vector with the initial values of U component when the color of the corresponding pixel is initially known or 0 otherwise; λ is a regularization parameter, and it is usually set to be 100 in our experiments.

Define a $MN \times MN$ matrix $\boldsymbol{L}$, whose (i, j)-th element is

$$\sum_{k \mid i,j \in w_k} \delta_{ij} - \frac{1}{n}\left(1 + \frac{1}{\frac{\varepsilon}{n} + \sigma_k^2}(Y_i - \mu_k)(Y_j - \mu_k)\right)$$

where δ_{ij} is the kronecker delta, μ_k and σ_k^2 are the mean and variance of the intensity component values of all pixels in window w_k.

The objective Eq. (6) can be represented to be

$$J(U) = U^T LU + \lambda(DU - U_0)^{\mathrm{T}}(DU - U_0) \tag{7}$$

It's obvious that there are two terms in our objective function of the U component values. The first term is the smoothness term based on the local linear relationship assumption. The second term is the data term based on the initial color constrains.

Different from the traditional colorization method, our smoothness term restricts the relationships between all of the pixel pairs within a local window rather than merely those of the center pixel and its neighboring pixels, and the initial color constrains are also relaxed in our method to obtain better estimation of the U component values, since in most cases the initial color is indicated roughly by the scribbles.

The optimal U component values can be obtained by solving the following sparse linear system:

$$(L + \lambda D)U = \lambda U_0 \tag{8}$$

The optimal V component values of each pixel can be computed similarly.

After estimating the U, V component values, the corresponding colorized image is output by holding the input grayscale image as Y component and converting them from YUV space back to RGB space.

3 Experiment results

In order to compare the performance of our colorization method with other colorization method objectively, experiments are carried out on various color images. Greyscale images, the input of colorization methods are converted from the original color images by

$$Y = 0.299R + 0.587G + 0.144B \tag{9}$$

Part of colorization results using test images extracted from standard image database (SIDBA) are presented in Fig. 2, in which (a) are the original color images, (b) are the grayscale images with seed pixels as an initial color condition, (c) are our colorization results and finally (d) are results of Levin[14]. In the experiment, the seed pixels are sampled uniformly from each original image according to certain ratio (we also call it the sampling ratio in the following description, and in Fig. 1 the sampling ratio is 0.3%). Clearly, under the same initial color conditions, our method gives better results, especially in preventing cross-object color mixture in the regions with complex structures.

(a) original color images (b) grayscale images with seed pixels (c) our results (d) results of Levin

Fig. 2 Colorization results using seed pixels as initial color conditions

It is worth mentioning that all colorization results of Levin in this paper are computed by the correlation based weighting function (as shown in Eq. (10) with exact solver, since the results shown in Ref. [12] were all obtained using it).

$$w_{pq} = 1 + \frac{1}{\sigma_p^2}(Y_p - \mu_p)(Y_q - \mu_p) \tag{10}$$

where Y_p is the intensity value of pixel p; μ_p and σ_p are the mean and variance of the intensities in a window around p.

The average estimation error of the U and V component values of colorization image at different sampling ratios are shown in Fig. 3 and Fig. 4 respectively. Obviously, when the sampling ratio is smaller, that is, less seed pixels are given, both colorization methods give higher estimation error; and when the sampling ratio increases, more seed pixels are given, the average estimation error of both method decreases gradually. But at all sampling ratios our colorization method gives smaller estimation error than those of Levin, which proves the effectiveness of our method.

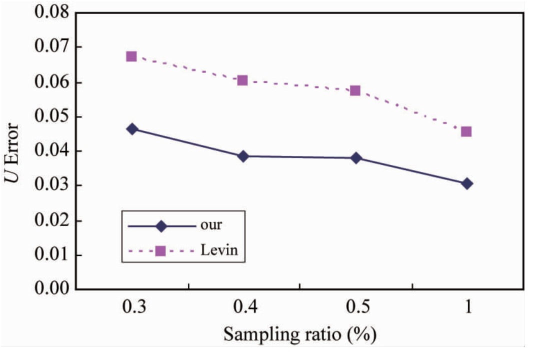


Fig. 3 Average absolute error of the U component values

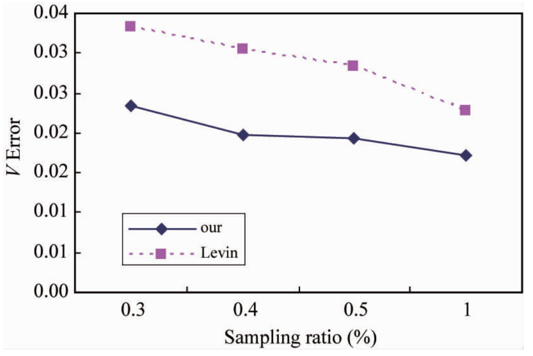


Fig. 4 Average absolute error of the V component values

The peak signal to noise ratio (PSNR) is also used in our experiments to verify the quality of the colorized images by both methods objectively. PSNR is a traditional measure widely used in image compression to estimate the quality of a reconstructed image compared with an original image.

PSNR in decibels (dB) is computed by

$$PSNR = 10 \log_{10} \frac{3MN \times 255^2}{\sum_{i=0}^{M} \sum_{j=0}^{N} |\tilde{I}(i, j) - I(i, j)|^2} \tag{11}$$

where M, N is the size of the input image; $I(i, j)$ is the real color vector $[R(i, j), G(i, j), B(i, j)]^T$; $\tilde{I}(i, j)$ is the color vector estimated by our method.

PSNR of both colorization results at different sample ratios are shown in Fig. 5. It can be found that PSNR of both colorization methods are improved with the increasing of the sampling ratio, while at all sampling ratios, PSNR of our method is higher than that of Levin. It proves that the colorization results of our method are more similar to the original image and have better quality under the same initial color conditions than the traditional method.

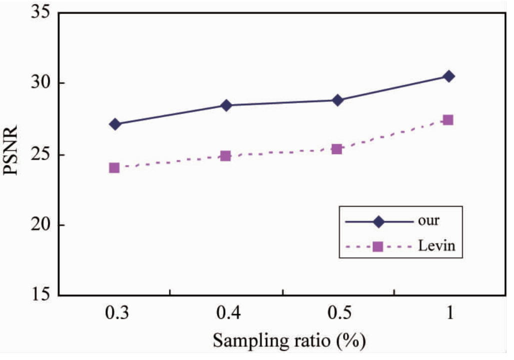


Fig. 5 PSNR of colorization results

Experiments using the user scribbles as initial color conditions are also carried out on various images, some results of which are presented in Fig. 6, in which (a) are the original color images, (b) are the grayscale images with the user scribbles as a initial color condition, (c) are our colorization results and (d) are the results of Levin. The colorized images comparison confirms that our method creates the result images with more vivid colors and sharper boundaries than the traditional method.

4 Conclusion

In this paper we introduce a grayscale image colorization method. It gives better colorization results than the traditional method. The aim of our colorization method is to minimize a quadratic objective function defined by the sum of the local costs based on a local linear relationship assumption under relaxed initial data constraints. The globally optimal values of two chromaticity components can be computed respectively by solving a sparse linear system of equations. As described above, our method enhances the PSNR of the colorization results, decreases the average estimation error objectively, and improves the subjective visual effect of the colorized images by making the image color more vivid and the region boundaries sharper than the traditional colorization method.

(a) original color images (b) grayscale images with scribbles (c) our results (d) results of Levin

Fig. 6 Colorization results using the user scribbles as initial color conditions

References

[1] He K, Sun J, Tang X. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011,33(12): 2341-2353

[2] Levin A, Lischinski D, Weiss Y. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(2): 228-242

[3] Markle W. The development and application of colorization. *SMPTE Journal*, 1984,93(7): 632-635

[4] Gonzalez R C, Woods R E. Digital Image Processing (3nd Edition). Beijing, Publishing House of Electronics Industry, 2011. 436-448

[5] Shi J, Malik J. Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905

[6] Sýkora D, Buriánek J, Žára J. Segmentation of black and white cartoons. In: Proceedings of Spring Conference on Computer Graphics, New York, USA, 2003: 245-254

[7] Welsh T, Ashikhmin M, Mueller K. Transferring color to greyscale images. *ACM Transactions on Graphics*, 2002, 21(3): 277-280

[8] Chen T, Wang Y, Schillings V, et al. Grayscale image matting and colorization. In: Proceedings of the Asian Conference on Computer Vision, Jeju Island, Korea, 2004. 1164-1169

[9] Lipowezky U. Grayscale aerial and space image colorization using texture classification. *Pattern Recognition Letters*, 2006, 27(4): 275-286

[10] Charpiat G, Hofmann M, Scholkopf B. Automatic image colorization via multimodal predictions. *Lecture Notes in Computer Science*, Volume 5304, 2008, 126-139

[11] Chia A, Zhuo S, Gupta R K. Semantic colorization with internet images. *ACM Transactions on Graphics*, 2011, 30(6), Article 156

[12] Horiuchi T. Estimation of color for gray-level image by probabilistic relaxation. In: Proceedings of the 16th International Conference on Pattern Recognition, Quebec City, Canada, 2002, vol. 3. 867-870

[13] Horiuchi T, Hirano S. Colorization algorithm for grayscale image by propagating seed pixels. In: Proceedings of the International Conference on Image Processing, Barcelona, Spain, 2003. 457-460

[14] Levin A, Lischinski D, Weiss Y. Colorization Using Optimization, *ACM Transactions on Graphics*, 2004,23(3): 689-694

[15] Noda H, Niimi M, Korekuni J. Simple and Efficient Colorization in YCbCr Color Space. In: Proceedings of the 18th International Conference on Pattern Recognition, Washington, USA, 2006. 685-688

[16] Sapiro G. Inpainting the colors, In: Proceedings of IEEE International Conference on Image Processing, Genoa, Italy, 2005. 698-701

[17] Yatziv L, Sapiro G. Fast image and video colorization using chrominance blending. *IEEE Trans. Image Processing*, 2006, 15(5): 1120-1129

[18] Lagodzinski P, Smolka B. Interactive Colorization Based on Hybrid Distance Transform. In: Proceedings of the IEEE International Conference on Computer Science and Automation Engineering, Shanghai, China, 2011. 538-543

Nie Dongdong, born in 1977. She is an associate professor in the College of Sciences, Yanshan University. Her research interests include image process, pattern recognition and computer vision.