IMAGE COLORIZATION VIA COLOR PROPAGATION AND RANK MINIMIZATION

Yonggen Ling, Oscar C. Au, Jiahao Pang, Jin Zeng, Yuan Yuan, Amin Zheng

The Hong Kong University of Science and Technology, Hong Kong
{ylingaa, eeau, jpang, jzengab, yyuanad, amzheng}@ust.hk

ABSTRACT

Image colorization aims to add colors to grayscale images, which used to be a time-consuming and tedious task that requires lots of human efforts. In this paper, we present a novel colorization method based on color propagation and rank minimization. Given a small portion of chrominance values and a grayscale image, we firstly propagate the known color values to other pixels to be colorized. As the colorized image after color propagation is not accurate, we then define a confidence matrix to measure the propagation fidelity. Finally, pixels that have propagated chrominance values with confidence are colorized by rank minimization, which exploits the redundancy of natural images. Experimental results on real data set show that our proposed method achieves state-of-the-art colorization quality.

Index Terms— Colorization, rank minimization, image restoration, matrix completion

1. INTRODUCTION

Due to the limitation of camera technology, photos and movies are all monochrome in the past. Colorizing those monochrome images perceptually to make them meaningful and visually appealing is of great interest. The key challenge of this problem is that there are a number of potential colors that can be assigned to each gray pixel. Thus this problem is highly under constrained and exist many solutions for the general case.

Recently, several effective algorithms have been proposed to reslove the colorization problem with reasonable amount of color cues as the input. These color cues can be categorized into three classes, namely, color scribbles, example images of similar colors and scattered chrominance values indicating desired colors of some pixels.

Information of color scribbles is usually delineated by manual labor. Colorization with this kind of input requires a user to mark color scribbles on the target image [1, 2, 3]. Besides color scribbles, in [4], scribbles that help to group the regions are also required, by which the number of color scribbles can be greatly reduced. All these methods assume that color image is locally smooth and propagate known colors inside scribbles to the neighbouring pixels to colorized based on an optimization framework.

Some methods require user to provide reference images that are similar to the image to be colorized [5, 6]. In [7], partial segmentation information are also needed. In this kind of methods, the key idea is to exploit multiple image features to transfer the color information from reference color images to the target gray image.

The last category of color cues is a small portion of scattered color labels [8, 9]. These given scattered chrominance values can be randomly or uniformly distributed in the image. Though scattered color labels can be regarded as “micro” scribbles, however, rather than marking manually, they are usually generated (or selected) by other algorithms. Therefore, in this work, we regard it as a different category of input color cue. Pang *et al.*’s work [8] learns a natural image dictionary using sparse representation from a set of natural images, and then, starting from the given chrominance values, colorizes the whole grayscale image patch by patch via sparse optimization. In [9], Wang *et al.* recover the unknown chrominance values by matrix completion. Some methods using this kind of color cue can also apply color scribbles, such as the approach of [9].

In this paper, we present a novel colorization method via color propagation and rank minimization using the last kind of color cues, namely, scatted color labels. We firstly propagate the known chrominance to grayscale pixels by the local texture and intensity similarity of natural image. Then a confidence matrix that capture the propagation accuracy is defined since the initial colorized image by color propagation is not accurate. Finally, we apply rank minimization to colorize the pixels that have chrominance values with confidence. Section 2 elaborates our proposed method. Experimental results are presented in Section 3. Finally, Section 4 concludes the paper.

2. COLORIZATION VIA COLOR PROPAGATION AND RANK MINIMIZATION

Our work is motivated by the recent development of matrix completion, which restores missing elements of a matrix from a portion of known values, and matrix recovery, which restores a matrix corrupted by noises and outliers. The most important spirit of matrix completion and recovery techniques

(a) (b) (c) (d)

Fig. 1. Colorization of a grayscale image with the proposed method. (a) Monochrome image with 0.5% scattered color labels. (b) Image after color propagation. (c) Confidence map of the propagated colors. (d) Colorization result.

is rank minimization. Since the problem of rank minimization is NP-hard, the nuclear norm (convex envelope of matrix rank) is widely used in the practice. Recently, Candes *et al.* [10, 11] show that low rank matrices can be recovered exactly from a small number of sampled elements under certain conditions. By virtue of this elegant property, rank minimization has been used in collaborative filtering [10], background modeling [11], image alignment, etc. To the best of our knowledge, there exist only one attempt [9] that applies matrix completion theory to image colorization. Based on different assumptions and inspiration, we apply rank minimization to image colorization in a different way.

Fig. 1 demonstrates the main process of our proposed method. Our approach is based on the intuition that natural image is NOT full of rank and color components on a natural image are highly correlated. Firstly, there are intrinsic similarity and redundancy on a natural image. Lots of image processing applications, such as de-noising, inpainting and super-resolution, exploit the redundancy in natural image and reach excellent performance in the literature. Secondly, dimension reduction algorithms, like principle component analysis (PCA), aim to find significant bases that can recover images with little distortion. The feasibility of dimensionality reduction techniques reveal the singular property of natural images. Generally speaking, if the rank of a natural image is full, its columns or rows are all linearly independent, and the colorization problem will be extremely hard to solve. Conversely, if the natural image is not full of rank, columns or rows are correlated, and the colorization problem becomes solvable. Lastly, the values of RGB color components are found to be highly correlated. They tend to gradually change in smooth region, and fast change in edge region simultaneously.

2.1. Notations

For better presentation, some notations are firstly introduced. For a matrix $\mathbf{X} \in \mathbb{R}^{m\times n}$, we denote $||\mathbf{X}||_0$ as the number of non-zero entries in $\mathbf{X}$ and $||\mathbf{X}||_1 = \sum_{i,j} |\mathbf{X}_{i,j}|$. $||\mathbf{X}||_F$ is the Frobenius norm (i.e. $||\mathbf{X}||_F = (\sum_{i,j} \mathbf{X}_{i,j}^2)^{1/2}$), $||\mathbf{X}||_*$ is the nuclear norm (i.e. $||\mathbf{X}||_* = \sum_{i=1}^{r} \sigma_i(\mathbf{X})$, $r = \min\{m, n\}$ and $\sigma_i(\mathbf{X})$ is the i^{th} eigen-values of $\mathbf{X}$). Moreover, let $\mathbf{I}_n$ be an identity matrix with size of $n \times n$, $\mathbf{X} \circ \mathbf{Y}$ be the Hadamard product of $\mathbf{X}$ and $\mathbf{Y}$ (i.e. $[\mathbf{X} \circ \mathbf{Y}]_{i,j} = [\mathbf{X}_{i,j}\mathbf{Y}_{i,j}]$). Matrix $\mathbf{X}$ is called indicator matrix if all the entries in $\mathbf{X}$ are either 0 or 1.

2.2. Problem Formulation

Let $\mathbf{Y} \in \mathbb{R}^{M\times N}$ be the grayscale natural image to be colorized, and $\mathbf{P} = [\mathbf{R}, \mathbf{G}, \mathbf{B}] \in \mathbb{R}^{M\times 3N}$ be the color image to be recovered with $\mathbf{R}$, $\mathbf{G}$ and $\mathbf{B} \in \mathbb{R}^{M\times N}$ being stacked horizontally, where $\mathbf{R}$, $\mathbf{G}$ and $\mathbf{B}$ are red, green and blue color components respectively. The definition of $\mathbf{P}$ that composes of $\mathbf{R}$, $\mathbf{G}$ and $\mathbf{B}$ being stacked horizontally helps to leverage the inter color correlation among the color components. Suppose $\mathbf{Y} = \alpha_1\mathbf{R} + \alpha_2\mathbf{G} + \alpha_3\mathbf{B}$. Two common monochrome image transforms in the literature are $\alpha_1 = \alpha_2 = \alpha_3 = \frac{1}{3}$ (average) and $\alpha_1 = 0.299$, $\alpha_2 = 0.587$, $\alpha_3 = 0.114$ (YUV). Let $\mathbf{P}' \in \mathbb{R}^{M\times 3N}$ be the initial colorized version of image $\mathbf{P}$ using color propagation (which will be elaborated in the next subsection), $\mathbf{N} \in \mathbb{R}^{M\times 3N}$ be a noise matrix, $\mathbf{W} \in \mathbb{R}_+^{M\times 3N}$ be the confidence matrix, whose elements represent the confidence of the initial colorized chrominance values. The colorization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{R},\mathbf{G},\mathbf{B},\mathbf{P},\mathbf{N},r} \quad & ||\mathbf{W} \circ \mathbf{N}||_F^2 \\ \text{subject to} \quad & \mathbf{P} = [\mathbf{R},\, \mathbf{G},\, \mathbf{B}] \\ & \mathbf{Y} = \alpha_1\mathbf{R} + \alpha_2\mathbf{G} + \alpha_3\mathbf{B} \\ & \mathbf{P}' = \mathbf{P} + \mathbf{N} \\ & \text{rank}(\mathbf{P}) \le r \end{aligned} \tag{1}$$

The second constrain can be rewritten as $\mathbf{Y} = \mathbf{PT}$, where

$$\mathbf{T} = [\alpha_1\mathbf{I}_N,\, \alpha_2\mathbf{I}_N,\, \alpha_3\mathbf{I}_N] \tag{2}$$

As the actual rank of $\mathbf{P}$ is unknown, we relax the fourth constrain by Lagrange multiplier λ, which serves as a balance factor. Therefore, (1) can be rewritten as

$$\begin{aligned} \min_{\mathbf{P},\mathbf{N}} \quad & ||\mathbf{W} \circ \mathbf{N}||_F^2 + \lambda \cdot \text{rank}(\mathbf{P}) \\ \text{subject to} \quad & \mathbf{Y} = \mathbf{PT} \\ & \mathbf{P}' = \mathbf{P} + \mathbf{N} \end{aligned} \tag{3}$$

In (3), the objective function contains two terms. The first term is the square of the Frobenius norm of element-wise weighted entries in the noise matrix $\mathbf{N}$, while the second term penalizes the rank of matrix $\mathbf{P}$, which leads to a singular solution. If the first term is replaced by l_0-norm and the weighted matrix $\mathbf{W}$ is an indicator matrix, the formulation (3) reduces to [9]. Since the regularized rank minimization in (3) is NP-hard, we relax the rank function with nuclear norm,

$$\begin{aligned} \min_{\mathbf{P},\mathbf{N}} \quad & ||\mathbf{W} \circ \mathbf{N}||_F^2 + \lambda ||\mathbf{P}||_* \\ \text{subject to} \quad & \mathbf{Y} = \mathbf{P}\mathbf{T} \\ & \mathbf{P}' = \mathbf{P} + \mathbf{N} \end{aligned} \tag{4}$$

The first constrain of (4) can be further relaxed by introducing the Lagrange multiplier η, and then we obtain a robust formulation,

$$\begin{aligned} \min_{\mathbf{P},\mathbf{N}} \quad & ||\mathbf{W} \circ \mathbf{N}||_F^2 + \lambda ||\mathbf{P}||_* + \frac{\eta}{2}||\mathbf{Y} - \mathbf{P}\mathbf{T}||_F^2 \\ \text{subject to} \quad & \mathbf{P}' = \mathbf{P} + \mathbf{N} \end{aligned} \tag{5}$$

If $\eta \to \infty$, (5) reduces to (4). With relaxation of (5), we can handle cases that there are potential outliers in the known monochrome image $\mathbf{Y}$ and noises in the given chrominance values.

2.3. Color Propagation

In this subsection, the process of computing the initial colorized image $\mathbf{P}'$ and the confidence matrix $\mathbf{W}$ that represents the confidence of the initial colorized chrominance values are introduced. Given a small portion of scattered chrominance values and a monochrome image, we propagate those known chrominance values to the unknown ones. The propagation procedure is pixel-based. Let p be the pixel whose color is unknown, q be the pixel whose color is going to be propagated to p. We define a confidence term $conf(p) \in (0, 1]$ for each pixel p, then

$$conf(p) = \begin{cases} w(p) & \text{if } w(p) \geq T_{prop} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$\mathbf{R}'(p) = \mathbf{R}(q^*),\ \mathbf{G}'(p) = \mathbf{G}(q^*),\ \mathbf{B}'(p) = \mathbf{B}(q^*) \tag{7}$$

$$w(p) = f(p, q^*),\ q* = \max_{q \in \Omega_p} f(p, q) \tag{8}$$

$$f(p, q) = \exp(-\frac{dist_t(p, q)}{\sigma_t^2}) \cdot \exp(-\frac{dist_i(p, q)}{\sigma_i^2}) \tag{9}$$

where Ω_p is the set of pixels whose chrominance values are known, $\mathbf{R}$, $\mathbf{G}$, $\mathbf{B}$ are the three color components in $\mathbf{P}$, $\mathbf{R}'$, $\mathbf{G}'$, $\mathbf{B}'$ are the three color components in $\mathbf{P}'$, $dist_t(p, q)$ is the Chi-square distance local texture feature distribution between p and q, $dist_i(p, q)$ is the Chi-square local intensity distribution distance between p and q, σ_t^2 and σ_i^2 are predefined constants. Texture feature distribution is calculated as follows:

i) We firstly apply Gabor filters with increment $\pi/8$ from 0 to $7\pi/8$ and five scales 0, 1, 2, 3, 4 to the whole grayscale image.

ii) Each pixel has 40-dimensional texture feature after previous step. Those features are then grouped by k-means clustering (k is set to be 128 empirically in this work). The cluster centers are taken as codewords, such that each pixel is associated with a codeword.

iii) The local texture feature distribution of a pixel is the distribution of codewords of a $(2K+1) \times (2K+1)$ patch whose center is the current pixel.

The intensity distribution is computed in a similar way, except that the input to the k-means clustering process is intensity instead of Gabor feature. Chi-square distance between two distribution $\mathbf{a}$ and $\mathbf{b}$ is defined as

$$dist(\mathbf{a}, \mathbf{b}) = 2 \sum_{i=1}^{n} \frac{|\mathbf{a}(i) - \mathbf{b}(i)|}{\mathbf{a}(i) + \mathbf{b}(i)} \tag{10}$$

where n is the dimension of distribution a and b.

To accelerate the searching process, for each pixel p to be colorized, we restrict Ω_p to be the set of pixels whose color are known and inside the window centered on p with size of $(2R+1) \times (2R+1)$.

For each pixel with known color, its confidence value is defined as 1 and its color components $\mathbf{R}'$, $\mathbf{G}'$, $\mathbf{B}'$ are the same as $\mathbf{R}$, $\mathbf{G}$, $\mathbf{B}$. After that, all the confidence values of the image together form a confidence matrix $\mathbf{W}' \in \mathbb{R}^{M \times N}$, while the propagated chrominance values form the initial color image $\mathbf{P}' \in \mathbb{R}^{M \times 3N}$. The confidence matrix $\mathbf{W}$ in the previous section is defined as

$$\mathbf{W} = [c\mathbf{W}'\, c\mathbf{W}'\, c\mathbf{W}'] \tag{11}$$

where c is a nonnegative real number that scales the confidence.

2.4. Solution via Iterative Convex Programming

Similar to [9], we introduce an auxiliary matrix $\mathbf{A}$ to decouple $\mathbf{P}$ in the objective function so as to solve problem 5. Therefore, problem 5 can be rewritten as

$$\begin{aligned} \min_{\mathbf{P},\mathbf{N},\mathbf{A}} \quad & ||\mathbf{W} \circ \mathbf{N}||_F^2 + \lambda ||\mathbf{P}||_* + \frac{\eta}{2}||\mathbf{Y} - \mathbf{A}\mathbf{T}||_F^2 \\ \text{subject to} \quad & \mathbf{P}' = \mathbf{P} + \mathbf{N} \\ & \mathbf{A} = \mathbf{P} \end{aligned} \tag{12}$$

The augmented Lagrange multiplier method in [12] is then applied to solve problem 12. The major difference of variable updating between [9] and the proposed method is the update of $\mathbf{N}$. In our work, the update of $\mathbf{N}$ in each iteration is

$$\forall i, j, \mathbf{N}_{i,j}^{k+1} = (\mathbf{\Theta}_{i,j}^k + \mu^k(\mathbf{P}'_{i,j} - \mathbf{P}_{i,j}^{k+1}))/(2\mathbf{W}_{i,j}^2 + \mu^k).$$

Fig. 2. 10 selected test images from Kodak PhotoCD.

3. EXPERIMENTAL RESULTS

In the section, we demonstrate the experimental results of our work and compare them with results of [1]. Fig. 2 shows 10 test images from Kodak PhotoCD [13]. These images are all of size 512×768 or 768×512. For each image, we randomly remove 99.9% of chrominance. For the parameters in our proposed method, we empirically set $\lambda = 10$, $\gamma = 100$ and $c = 1$. Though these parameters may not be optimal, fortunately, from the experiments that have been done, our proposed method is not sensitive to the parameters setting in a wide range. To quantitatively assess the performance of different colorization algorithms, we adopt color peak signal-to-noise ratio (CPSNR) as an objective measurement, which is defined as follows:

$$\text{CPSNR} = 10\log_{10}\frac{255^2}{\text{MSE}} \tag{13}$$

$$\text{MSE} = \frac{1}{3MN}\sum_{\mathbf{C}=\{\mathbf{R},\mathbf{G},\mathbf{B}\}}\sum_{i=1}^{M}\sum_{j=1}^{N}(\mathbf{C}(i,j)-\mathbf{C}'(i,j))^2 \tag{14}$$

where $\mathbf{R}$, $\mathbf{G}$, $\mathbf{B}$ are RGB components of the ground-truth image, and $\mathbf{R}'$, $\mathbf{G}'$, $\mathbf{B}'$ are RGB components of the recovered image by colorization algorithm with grayscale image and a small portion of scattered chrominance values as input.

Numerical results are tabulated in Table 1. It revels that, with scattered color cues, our proposed method achieves superior results in terms of CPSNR.

Image	Method in [1]	Proposed method
1	30.10	**30.61**
2	27.94	**28.97**
3	28.68	**29.20**
4	28.18	**28.45**
5	**31.91**	31.84
6	28.36	**29.25**
7	33.08	**33.75**
8	31.98	**33.40**
9	**27.55**	27.49
10	29.00	**30.17**

Table 1. Comparison of CPSNR between different colorization methods.

(a) (b) (c)

Fig. 3. Comparison between different methods with test image 1. (a) Input grayscale image and 0.1% chrominance values. (b) Result of the proposed method. (c) Result of the method in [1].

Fig. 3 presents a visual comparison of different colorization results. Although both two methods can recover the color given sufficient labels, a close inspection at the recovered images reveals differences in intensity, chrominance details. For the results obtained by [1], the colorized image looks blurred and unnatural. This is because Levin *et al.*'s work is based on the assumption that neighbouring pixels with similar intensity should have similar colors, which does not always correct. For the results computed by our proposed method, the recovered image looks better.

Notice that the proposed colorization method can be applied to video colorization, too. In this situation, all the video frames are firstly separated into different groups. Then frames in the same group are stacked horizontally to optimize as a whole so that the inter correlation between them can be exploited. Finally, colorization is performed to the stacked frames. For other applications of the proposed method, one of them is lossy image or video compression. For the encoder, the grayscale image and a small portion of chrominance values are encoded. For the decoder, the encoded information is decoded and the image or video is then recovered through colorization.

4. CONCLUSIONS

In this paper, we focus on the problem of image colorization which adds color onto a grayscale image. We tackle the colorization problem by color propagation and rank minimization. We firstly propagate the input scattered chrominance values to the grayscale pixels to be colorized based on local similarity of natural image. By leveraging the local similarity on natural images, the advantages of texture similarity and intensity similarity can be combined. We then applies the rank minimization framework that further exploits the redundancy of natural image. Numerical and visual comparisons show that our proposed method leads to colorization results of high-quality.

5. REFERENCES

[1] A. Levin, D. Lischinski, D. and Y. Weiss, "Colorization using optimization," in *Proc. ACM SIGGRAPH Conf.*,

vol. 23, pp. 689-694, 2004.

[2] L. Yatziv ang G. Sapiro, "Fast image and video colorization using chrominance blending," *IEEE Trans. on Image Processing*, vol. 15, pp. 1120-1129, 2006.

[3] X. Chen, D. Zou, Q. Zhao and P. Tan, "Manifold preserving edit propagation," *IEEE Trans. on Graphics*, vol. 31, 2012.

[4] Q. Luan, F. Wen, D. Cohen-Or, L. Liang, Y.-Q. Xu, and H.-Y. Shum, "Natural image colorization," in *Proc. 18th Eurograph. Symp. Rendering*, pp. 309-320, 2007.

[5] R. Gupta, Y.-S. Chia, D. Rajan, E. Ng and Z. Huang, "Image colorization using similar images," in *Proc. 20th ACM Int. Conf. on Multimedia*, 2012.

[6] X. Liu, L. Wan, Y. Qu, T.-T. Wong, S. Lin, C.-S. Leung, and P.-A. Heng, "Intrinsic colorization," *ACM Trans. on Graphics*, vol. 27, pp. 152:1-152:9, 2008.

[7] R. Irony, D. Cohen-Or, and D. Lischinski, "Colorization by example," in *Proc. 16th Eurograph. Symp. Rendering*, pp. 201-210, 2005.

[8] J. Pang, O. C. Au, K. Tang and Y. Guo, "Image colorization using sparse representation," In *Proc. 38th Int. Conf. Acoust., Speech Signal Processing*, 2013.

[9] S. Wang and Z. Zhang, "Colorization by matrix completion," in *Proc. Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[10] E. J. Cands and B. Recht. "Exact matrix completion via convex optimization," *Found. of Comput. Math.*, pp. 9 717-772, 2008.

[11] E. J. Candes, X. Li, Y. Ma, and J. Wright. "Robust principal component analysis?" *Journal of ACM 58(1)*, pp. 1-37, 2009.

[12] Z. Lin, M. Chen and L. Wu, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *UIUC Technical Report UILU-ENG-09-2215*, 2009.

[13] R. Franzen, "Kodak lossless true color image suite," `http://r0k.us/graphics/kodak/`, 2012.