Structure-Preserving Colorization Based on Quaternionic Phase Reconstruction

Zhou Yu, Yi Xu, Xiaokang Yang, and Li Song

Shanghai Key Laboratory of Digital Media Processing and Transmissions
Institute of Image Communication and Information Processing
{zhouyuyz,xuyi,xkyang,song_li}@sjtu.edu.cn

Abstract. A novel semiautomatic colorization method is proposed based on quaternionic phase reconstruction. In this method, each color pixel is represented as a quaternion, whose polar magnitude and polar phase are recovered from the intensity of original grayscale image and color scribbles of user's manual input, respectively. To conduct structure-preserving colorization, color diffusion is restrained across global image structures, which are extracted using hierarchical edge representation along with structural importance measurement. To identify local spatial relationship between neighboring pixels, Gabor wavelets are applied to compute the similarity of local phase patterns. Our method is highlighted in well preserving image structures during colorization, where the color image is acquired by solving a linearly constrained quadratic optimization problem. Specifically, we develop a method to guide the user to scribble on the monochrome image, so that effective color propagation from less manual input can be expected. Experimental results demonstrate that our colorization method outperforms the state-of-the-art method in structural preservation and relatively better colorization results are available if the proposed rule of scribble user guidance is adopted.

Keywords: Colorization, structure preserving, scribble user guidance.

1 Introduction

Colorization enhances visual appearance of traditional black and white movies by adding color without large amount of manual work. Besides that old monochrome images can achieve better perceptual visual effects after colorization, image re-colorization is also an interesting topic for photo editing software.

Major problems involved in colorization include high computational cost and time-consuming operation. It is unacceptable if we interactively colorize an image with a long time to wait. To simplify the colorization process, Levin[1] proposed an interactive video colorization technique without precise image segmentation and accurate region tracking. His method depends on an assumption that nearby pixels with similar intensity should have similar chromatic values. However, this assumption only considers the intensity distribution of image, which is sensitive to noises and illumination variations. No image structure information is taken into account. Oversmoothness between two regions can be observed with some blocks wrongly colorized.

P. Muneesawang et al. (Eds.): PCM 2009, LNCS 5879, pp. 849–859, 2009.

Another work presented by Liron Yatziv[2] formulates the colorization problem as a minimization process of a cost function. The propagation of color is conducted on the shortest integral gradient path. This method needs to find such a shortest path around each uncolored pixel, resulting in computational consuming operation. Manga colorization proposed by Qu[3] propagates colors over regions which exhibit pattern-continuity as well as intensity-continuity. This method extracts pattern vectors via multi-scale and multi-orientational Gabor filtering. Structural importance of these patterns is not measured. Hence, colorization around local textures and global structures are equally treated. Welsh[4] transfers the color from reference color image to target monochrome image. The visual quality of colorized image completely relies on the similarity of two images' local luminance distribution and textural information. Irony[5] colorizes grayscale images by transferring color from a segmented example image. Rather than rely on a series of independent pixel-level decisions, this method accounts for the higher-level context of each pixel. Similar to Welsh's work, the colorization quality is dependent on the proper selection of reference blocks within the example image.

In this paper, a new colorization system is advanced to pay much attention to the structure distribution of original grayscale image and scribbled regions. Scribble user guidance during interactive operations is also considered to reduce the manual input. The proposed system conducts colorization process by solving quaternionic phase reconstruction problem. Each color pixel is represented as a quaternion, whose polar magnitude and polar phase are recovered from the intensity of original grayscale image and color scribbles of user's manual input, respectively. The global structures of the original grayscale image are detected in scale space based on structural importance measurement using hierarchical edge representation. Color diffusion is restrained across these significant structures which always appear as the object contours. Upon the fact that phases encode the local image structures, only the image elements with similar Gabor phase patterns contribute to the color propagation around each pixel. Specifically, we develop a method to guide the user to scribble on the monochrome image, so that effective color propagation from less manual input can be expected.

The rest of this paper is organized as follows. In Section 2, specifications of the proposed colorization method are provided. This is followed by the discussion of rule design for scribble user guidance in Section 3. Section 4 gives our colorization results and experimental comparison with the state-of-the-art methods. Conclusion remarks are drawn in Section 5.

2 Colorization Based on Quaternionic Phase Reconstruction

Before user's scribbling, our colorization system preprocesses the monochrome image and offers user some guidance about where to scribble. Similar to other semi-automatic colorization methods, users can iteratively input scribbles until satisfying about the colorization results.

2.1 Quaternionic Phase Feature Space

Most colorization methods are implemented in YUV color space. In this section, we conduct colorization in polar quaternion space. Each color pixel is represented as a pure quaternion $q = a + bi + cj + dk$, where a is zero, b , c , d represent r, g, b channel respectively. In polar coordinates, a quaternion can be formed as,

$$q = |q| e^{i\phi} e^{k\psi} e^{j\theta} \tag{1}$$

where ϕ , ψ and θ are three quaternion phases, which can be computed as

$$\psi = -\frac{\arcsin(2(bc - ad))}{2} \tag{2}$$

$$\begin{cases} \theta = \frac{1}{2}\arctan 2(2(bd + ac), a^2 + b^2 - c^2 - d^2) \\ \phi = \frac{1}{2}\arctan 2(2(cd + ab), a^2 - b^2 + c^2 - d^2) \end{cases}_{\cos 2\psi \neq 0} \tag{3}$$

Quaternion phase ϕ should be modified if $q = -|q| e^{i\phi} e^{k\psi} e^{j\theta}$

$$\begin{cases} \phi = \phi + \pi, & if\ \phi < 0 \\ \phi = \phi - \pi, & otherwise \end{cases} \tag{4}$$

The use of quaternion phase to represent color has several advantages. In general, there is more or less interrelationship between the chromatic channels in the commonly-used color spaces, e.g. RGB space. However, there is little correlation information between quaternion phases ϕ , ψ and θ . As for YUV color space, which is known for the independence between the channels, uneven illumination would induce side effects to UV channels. Obviously, it has no impact on quaternion phases of color image, as referred to (1). Hence, it is more reasonable to treat quaternion phases separately and conduct colorization based on quaternionic phase reconstruction.

2.2 Global Image Structure Extraction via Significance Measurement

Global image structures always appear as object contours. In colorization process, color diffusion should be suppressed across these contours to avoid color leakage between different objects. In this section, we extract global image structures in scale space using hierarchical edge representation along with structural importance measurement.

Hierarchical edge representation is produced using multi-scale Canny edge detectors, in which image is convolved with a set of Gaussian filters and edges are detected based on the first derivative of the Gaussian filtering results. The Gaussian filters are designed with different scales. With a small Gaussian filter support, the resulting edge map would include many fine details of image. With a large Gaussian filter support, many fine edges disappear; just blurry edges will be remained. It is noted that global image structures can capture our eyes all the time when we go near from a far distance. That is, significant global image contours should have a long lifetime in the scale space. Here

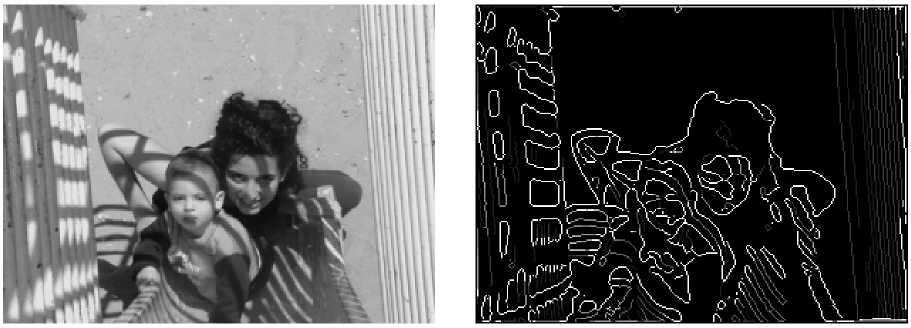


Fig. 1. Edges with different significance (global structures with long life-time in scale space are shown in high intensity and vice versa)

we introduce the work of Alexandrina Orzan[6] to measure the significance of image structures.

Given hierarchical edge representation produced using multi-scale Canny edge detectors, we assign significance measure to each edge point as follow:

1. We extract two edge maps $p(\sigma_{current})$ and $p(\sigma_{current} - \Delta(\sigma))$ using two specified Canny edge detectors, which respectively have the Gaussian deviations of $\sigma_{current}$ and $\sigma_{current} - \Delta(\sigma)$. The parameter $\sigma_{current}$ is initialized as $\sigma_{\max}$.
2. For those points belonging to $p(\sigma_{current}) \cap p(\sigma_{current} - \Delta(\sigma))$, we assign significance value as $e(p(\sigma_{current} - \Delta(\sigma))) = (\sigma_{\max} - \sigma_{current} + \Delta(\sigma))/\Delta(\sigma)$.
3. For points of $p(\sigma_{current}) / p(\sigma_{current} - \Delta(\sigma))$, we assign their significance as $e(p(\sigma_{current} - \Delta(\sigma))) = (\sigma_{\max} - \sigma_{current})/\Delta(\sigma)$, where $\sigma_{\max}$ is set as 4 and $\Delta(\sigma)$ is 0.4.
4. Reduce $\sigma_{current}$ by $\Delta(\sigma)$ and repeat the process from step 1 until $\sigma_{current}$ becomes 1.

During the iteration, the significance value of edge point is calculated. An example is presented in Fig.1 to show image structures with different significance values.

2.3 Smoothness Constraint and Structure Preservation for Colorization

Since colorization is an ill-posed problem, it is important to impose smoothness constraint on the color propagation within the homogeneous object components. Thus it is a key issue to define homogeneity. Levin[1] proposed to find homogeneous image elements depending on the intensity similarity. In Qu's work[3], homogeneous image elements are considered to have similar local patterns, which are depicted as a multi-dimensional feature vector obtained from Gabor filtering. Rather than using Gabor wavelet coefficient, we only employ Gabor phases to establish homogeneity metric. Phases encode spatial relation information of image and maintain stability under noises and illumination variations.

Since phase patterns can be used to identify local image structures, we establish a metric as (5) to check if two pixels $\mathbf{x_i}, \mathbf{x_j}$ exist in a homogeneous neighborhood, where $I(\bullet)$ is the intensity function of original grayscale image, operator $*$ denotes convolution operator and arg[.] extracts phase component from the convolution result. As shown in (5)-(7), a set of Gabor functions $g(\bullet)$ are designed to extract the local phase patterns across scale space as well as along multiple orientations. When two pixels $\mathbf{x_i}, \mathbf{x_j}$ belong to the same homogeneous neighborhood, the phase difference $\mathrm{H}(\mathbf{x}_i, \mathbf{x}_j)$ would have a small value.

$$\mathrm{H}(\mathbf{x}_i, \mathbf{x}_j) = \sum_{\sigma \subset \mathbf{S}} \sum_{\theta \subset \mathbf{\theta}} \left[\arg\left(I(x_i, y_i) * g_{\sigma,\theta}(x_i, y_i)\right) - \arg\left(I(x_j, y_j) * g_{\sigma,\theta}(x_j, y_j)\right) \right] \tag{5}$$

$$g(x, y; \sigma) = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \exp\left(\mathrm{i} \cdot 2\pi \frac{x'}{\lambda}\right),\ \mathrm{i}^2 = -1 \tag{6}$$

$$x' = x\cos\theta + y\sin\theta,\ y' = -x\cos\theta + y\sin\theta \tag{7}$$

Inspired by Levin's method, we formulate colorization problem by minimizing the following cost function, as shown in (8). It can be treated as a linearly constrained optimization problem,

$$E(\Phi) = \sum_{\mathbf{p}} \left(\Phi(\mathbf{p}) - \frac{\sum_{\mathbf{q} \subset N(\mathbf{p})} W_{\mathbf{pq}} \Phi(\mathbf{q})}{\sum_{\mathbf{q} \subset N(\mathbf{p})} W_{\mathbf{pq}}} \right)^2,\ \Phi = \{\phi, \psi, \theta\} \tag{8}$$

$$W_{\mathbf{pq}} = W^e_{\mathbf{pq}} W^g_{\mathbf{pq}} W^i_{\mathbf{pq}} \tag{9}$$

$$W^e_{\mathbf{pq}} = \exp(-|e(\mathbf{p}) - e(\mathbf{q})|^2) \tag{10}$$

$$W^g_{\mathbf{pq}} = \exp(-\mathrm{H}(\mathbf{p}, \mathbf{q})) \tag{11}$$

$$W^i_{\mathbf{pq}} = \exp(-(I(\mathbf{p}) - I(\mathbf{q}))^2 / 2\sigma^2_{\mathbf{p}} \tag{12}$$

where $N(\bullet)$ is a 3*3 neighborhood system and $W_{\mathbf{pq}}$ imposes a smoothness constraint for the color propagation between pixel **p** and **q**. To avoid color leakage across global structures, W^e_{pq} is used to check whether there is a significant object contour between pixel **p** and pixel **q**. The term W^g_{pq} enforces color propagation within structure continuous regions. Another complementary term W^i_{pq} is used to enforce color propagation within intensity continuous regions, which can help to alleviate phase singularity problem [7]. Parameter $\sigma_{\mathbf{p}}$ is the intensity variance around pixel **p**.

2.4 Quaternionic Phase Reconstruction

Once the cost function (8) is established, we can use general tools in matlab to conduct this minimization process. The three quaternion phases are individually reconstructed by adopting linearly constrained quadratic optimization and corresponding phase information provided by user's color scribbles. Fig.2 presents an illustration of two produced color images, using our method and Levin's method[1], respectively. It can be observed that our method achieved better colorization results across object contours.

(a)

(b)

Fig. 2. Comparison of our method with Levin's (a) our result (b) Levin's result

3 User Scribble Guidance Module

As for some images with complex structures, effective colorization needs large amount of careful scribbles. It would be impressive for users if the interactive colorization method can provide good visual quality with rather small amount of scribbles. Indeed, where and how to scribble is a troublesome problem for users. However, this problem is ignored in the state-of-the-art works. In this section, an attempt is conducted to provide some useful guidance for users where to scribble. These locations automatically provided by our approach help the user to achieve good colorization result with relatively small amount of manual input.

The scribble user guidance is based on a simple premise: the locations where the image elements result in better reconstruction of original grayscale image will leads to better colorization in the unsettled chromatic image. The curves covering these locations

Table 1. Pseudocode of our user guidance module

1. Start with calculating intensity correlation weight matrix

$$W_{pq}^{i} = \exp(-(I(p) - I(q))^2 / 2\sigma_p^2 \tag{13}$$

2. Initialize error image I_{err} based on (14):

$$I_{err} = \text{abs}\left(W^{i} I - I\right), \tag{14}$$

where operator $\text{abs}(\bullet)$ computes the absolute value.

3. Segment grayscale image into disconnected regions using canny edges of the coarsest scale σ_0 ($\sigma_0 = \sigma_{\max} = 4$); Here we use $\mathbf{C_T}$ to denote the canny edge set, T=0.

4. For each segmented component, locate the top 10 points with the biggest errors in image I_{err} , fitting these points as a spline which is thick enough to cover all the points. Thus, we get the initial scribble set $\mathbf{M_0}$ and enter into the following iterative refinement of scribble set.

5. For iteration T, reconstruct the grayscale image I_T based on (15) using current scribbles $\mathbf{M_T}$,

$$E(I) = \sum_{p \subset \mathbf{M}} \left(I(\mathbf{p}) - \frac{\sum_{\mathbf{q} \subset N(\mathbf{p})} W_{\mathbf{pq}}^{i} W_{\mathbf{pq},T}^{e} I(\mathbf{q})}{\sum_{\mathbf{q} \subset N(\mathbf{p})} W_{\mathbf{pq}}^{i} W_{\mathbf{pq},T}^{e}} \right)^2 , \tag{15}$$

$$W_{\mathbf{pq},T}^{e} = \exp(-\left|e_T(\mathbf{p}) - e_T(\mathbf{q})\right|^2) \ , \tag{16}$$

$$e_0(\mathbf{p}) = \begin{cases} 1, & \text{if } \mathbf{p} \subset \mathbf{C}_0 \\ 0, & otherwise \end{cases} \tag{17}$$

6. Compute error image $I_{err} = abs\left(I_T(i) - I\right)$, extracting those pixels $\mathbf{x}$ which satisfies the constraint $I_{err}(\mathbf{x}) > 0.05$ and thus forming a pixel set $\mathbf{S}$.

7. Segment $\mathbf{S}$ into disconnected regions using canny edges of scale $\sigma_T = \sigma_{T-1} - \Delta(\sigma)$, extract new canny edge set $\mathbf{C_T}$ and assign new significance to it. Please read Section 2.3 to get detail information.

8. For each segmented element, locate the top 10 points with the biggest errors in image I_{err} , fitting these points as a spline which is thick enough to cover all the points. Thus, we get the refined scribble set $\mathbf{M_T}$.

9. Continue the iteration from 5 until the current error image I_{err} has less than 10% of the total pixels larger than threshold 0.05.

will provide the user a good guidance where he should scribble. The pseudocode of the proposed user guidance module is described in Table 1.

Here Levin's minimization method [1] is imposed on the cost function (15).

Generally, such an iterative process will finish after two or three iterations, which takes just a little extra computational cost. Moreover, these scribbles are generated automatically. We can design an interactive color selection process for these scribbles, having no need to retrace them carefully with a color pen. To achieve colorization performance comparable to Levin's work presented on the website http://www.cs.huji.ac.il/~yweiss/Colorization/, our method needs much less guidance scribbles, as illustrated in Fig. 3.

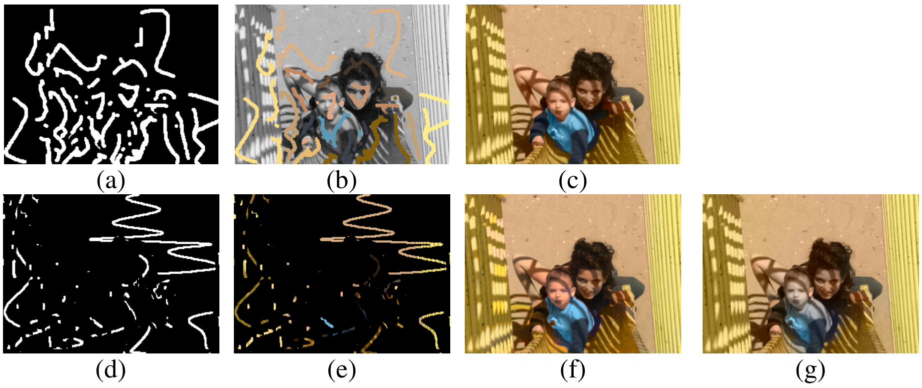
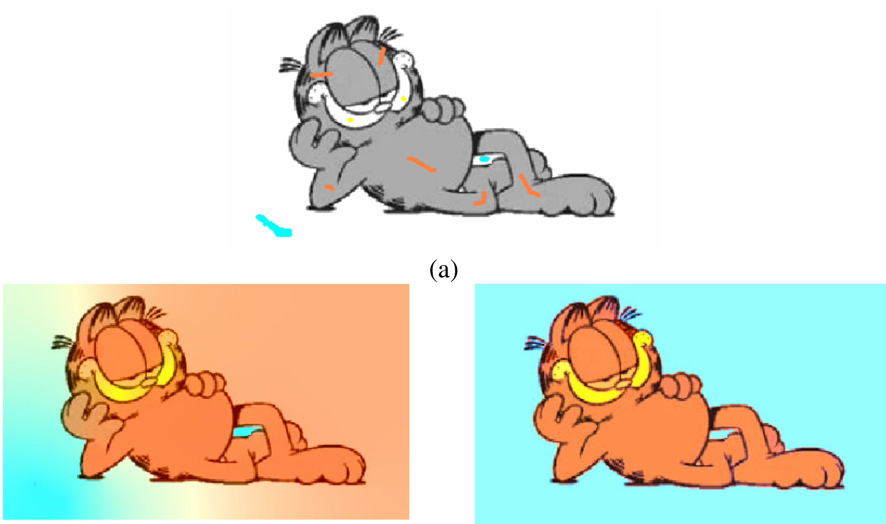


Fig. 3. Results of our user guidance system as compared with Levin's work [1]. (a) (b) Levin's scribbles (c) Levin's result (d) (e) scribbles automatically extracted by our scribble guidance module (f) results using our colorization method based on guidance scribbles (g) result using Levin's method based on guidance scribbles.

4 Experimental Results

We testify our colorization system on many images and compare it with Levin's method, which is considered as the fundamental block of the state-of-the-art colorization works. It can be observed from Fig. 4, Levin's method fails to colorize the background of Garfield. Moreover, our method results in less color leakage across the contours of Garfield. Fig. 5 and Fig. 6 show some colorization results of our method. Our guidance system is also tested using various images. Fig. 3, Fig. 7 and Fig. 8 illustrate that guiding scribbles automatically generated by our method can produce a color image with good visual quality. This tells us the generated scribbles should encode the major image pattern information. In addition, our method needs much less guidance scribbles to obtain comparable colorization performance to Levin's work. If we apply such few scribbles to Levin's method, the results become worse. From Fig. 3, Fig. 7 and Fig. 8, we can see when structure of image is complex and color of image is diverse, our method show more advantages with much less manual label than Levin's work. Otherwise, our method show little improvement, but still provides users useful guidance for scribble generation. All the experimental results validate the effectiveness of our guidance system as well as our colorization method.

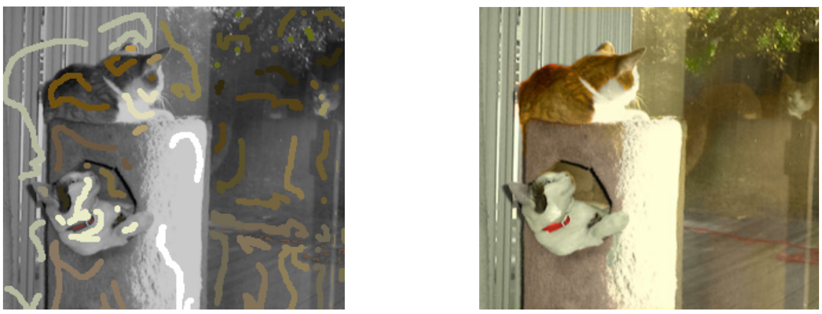


(a) (b) (c)

Fig. 4. Colorization of Garfield. (a) scribbled image (b) Levin's result (c) our result.

(a) (b)

Fig. 5. Colorizing still image based on our colorization method



(a) (b)

Fig. 6. Colorizing still image using our colorization method

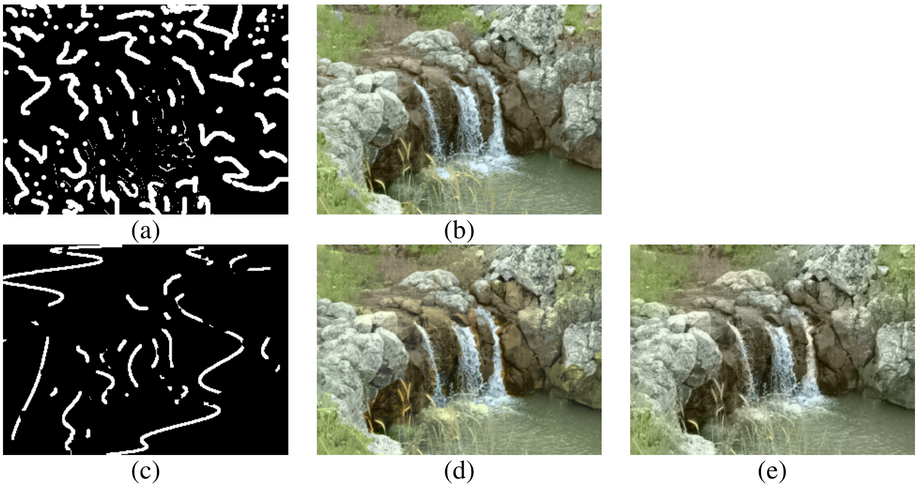


Fig. 7. Results of user scribble guidance system as compared with Levin's work [1]. (a) Levin's scribbles (b) colorization result using Levin's method with Levin's scribbles (c) scribbles extracted by our user guidance system (d) colorization result using our method based on (c)'s scribbles (e) colorization result using Levin's method based on (c)'s scribbles.

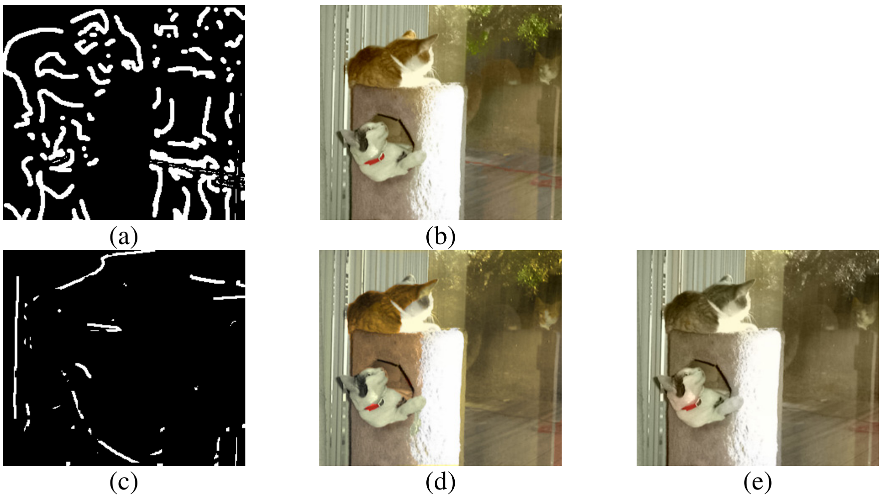


Fig. 8. Results of user scribble guidance system as compared with Levin's work [1]. (a) Levin's scribbles (b) colorization result using Levin's method with Levin's scribbles (c) scribbles extracted by our user guidance system (d) colorization result using our method based on (c)'s scribbles (e) colorization result using Levin's method based on (c)'s scribbles.

5 Conclusion

In this paper, a new system to colorize monochrome image is proposed based on quaternionic phase reconstruction. It preserves image structures of the original monochrome image and the manually scribbled regions during color diffusion, so that better colorization result can be expected. It also presents a guidance system to help the user scribble on monochrome images. This guidance system frees the user from roundly testing and time consuming colorization process. Extensive experimental results demonstrate that our novel colorization system is efficient to colorize most monochrome images with greatly reduced manual inputs.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China (60702044 and 60828001), Research Fund for the Doctoral Program of Higher Education of China (200802481006), NCET-06-0409, Cultivation Fund of the Key Scientific and Technical Innovation Project of MOE (706022), and the 111 Project (B07022).

References

1. Levin, A., Lischinski, D., Weiss, Y.: Colorization using optimization. In: ACM SIGGRAPH, vol. 3, pp. 689–694 (2004)
2. Yatziv, L., Sapiro, G.: Fast image and video colorization using chrominance blending. IEEE Transactions on Image Processing 15(5), 1120–1129 (2006)
3. Qu, Y.G., Wong, T.T., Heng, P.A.: "Manga Colorization". ACM Transactions on Graphics 25(3), 1214–1220
4. Welsh, T., Ashikhmin, M., Mueller, K.: Transferring color to grayscale images. ACM Transaction on Graphics 21(3), 277–280 (2002)
5. Irony, R., Cohen-or, D., Lischinski, D.: Colorization by example. In: Proc. of Eurographics Symposium on Rendering, pp. 201–210 (2005)
6. Orzan, A., Bousseau, A., Barla, P., Thollot, J.: Structure-preserving manipulation of photographs. In: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering, pp. 103–110 (2007)
7. Zhou, J., Xu, Y., Yang, X.K.: Quaternion wavelet phase based stereo matching for uncalibrated images. Pattern Recognition Letters 28(12), 1509–1522 (2007)