Image composition with color harmonization

Congde Wang (王从德)*, Rong Zhang (张 荣), and Fan Deng (邓 璠)

Department of Electronic Engineering and Information Science,
University of Science and Technology of China, Hefei 230027, China
**E-mail: hwlongw@mail.ustc.edu.cn*

Received September 1, 2008

Image matting and color transfer are combined to achieve image composition. Firstly, digital matting is used to pull out the region of interest. Secondly, taking color harmonization into account, color transfer techniques are introduced in pasting the region onto the target image. Experimental results show that the proposed approach generates visually pleasing composite images.

OCIS codes: 100.2960, 100.2980, 330.1720, 100.2000.
doi: 10.3788/COL20090706.0483.

The technology of image composition allows various image elements to be mixed in a new image. Producing a realistic composite image is widely considered to be very important for computer graphics applications. In general, image composition process consists of two steps. First, the region of interest is extracted from a foreground image. Second, the region is pasted onto a background image (target image) seamlessly, without bringing any visual artifacts. The difficulty of the first step lies in extracting the concerned object as accurately as possible from an arbitrary image, while consuming minimal time/interaction of user. In the second step, the luminance and color of the concerned object have to be changed probably according to the background image.

Several image segmentation and composition algorithms capable of producing high-quality results have been developed in computer vision community[1−5]. Intelligent scissors was used to search the object boundary between user clicks by finding the shortest path on a weighted graph[1]. Then the extracted objects could be scaled and rotated, and then composited directly with a background image. Bayesian estimation was applied for automatic image segmentation[2]. Grabcut provided a foreground object segmentation solution via graph cut, and pasted the object onto a background with a similar way as intelligent scissors[3]. Grabcut and Poisson image editing[6] were combined together and improved in drag-and-drop pasting[4] method to achieve robust and visually pleasing results. A hybrid algorithm[5] based on seeded region growing and k-means clustering was proposed to improve image object segmentation result.

Generally, the methods mentioned above have two shortcomings. Firstly, these methods extract foreground objects in a "hard" manner[1−3,5], which means that a pixel is labeled to be either foreground or background. However, values of pixels on a foreground object border vary gradually from the object to background and the translucent areas might exist, due to which a single pixel may be a combination of foreground and background. Therefore "hard" labeling here is incorrect. The technology of digital image matting[7,8], which is used to estimate an opacity (alpha) value for each pixel as well as the foreground color, is introduced in this letter to fix this problem.

Secondly, composition of extracted objects with background image draws less attention to researchers than segmentation in the first step. In some previous work, the object was just pasted onto target image directly without adjusting its chroma and intensity[1,9]. This method is prone to producing visually inauthentic images with obvious artificial areas, as a result of different lighting environment applying on foreground and background, respectively. Sometimes experienced artists can manually tune parameters to harmonize the result image. But that is really complicated and automating the painstaking process would be very profitable. Poisson method is used to merge the foreground and background in drag-and-drop pasting. But it changes foreground color in an uncontrollable manner, and produces unacceptable results sometimes. Color transfer is the process of adjusting an image's color to consist with that of another image[10,11], which is applied in many fields[12,13]. Natural shadow matting uses color transfer to remove shadow in digital images[13]. It changes the color characteristics of the pixels in the shadow to accord with the ones around the shadow. Inspired by this, we apply color transfer to solve the color harmonization problem between foreground objects and background. The color of extracted object is transferred according to the background color.

As mentioned above, in this letter, digital matting and color transfer are combined to form a complete composition system. Specifically, Bayesian matting[7], one of the digital matting approaches, is introduced here to "softly" extract the object of interest from a foreground image, and then to paste the object onto a background image. At the same time, Neumann's color transfer method[8] is applied to harmonize the color of foreground object and background. Both two parts are briefly described in the following two paragraphs, respectively.

Bayesian matting is one of the early methods that can produce high-quality results in many cases. An image can be represented by the compositing equation:

$$C_i = \alpha F_i + (1-\alpha)B_i, \tag{1}$$

where i = 1, 2, 3 represent three channels of the image, C_i, F_i, and B_i are the pixel's composite, foreground, and background colors, respectively, and α is the pixel's opacity component used to linearly blend between foreground and background. The problem is under-constrained since F_i, B_i, and α are all unknown. The Bayesian matting assumes that the input image has already been segmented into three regions: "background", "foreground", and "unknown", while foreground and background regions have been delineated manually and conservatively, generating a trimap. For each pixel in the unknown region, Bayesian matting builds the color probability distribution using the known and previously estimated foreground and background colors within the pixel's neighborhood. Maximum a posterior (MAP) estimation is then used to calculate F_i, B_i, and α. After every α value of each pixel in the image is computed, the object of interest is extracted softly.

Neumann's method transfers the color characteristic of a style image into a target image[10]. This technique uses a permissive, or optionally strict, three-dimensional (3D) histogram matching, which is similar to the sequential chain of conditional probability density functions. Given two 3D histograms, ${}^{(1)}P_{\mathrm{p}}(x, y, z)$ for a style image and ${}^{(2)}P_{\mathrm{p}}(x, y, z)$ for a target image, Neumann's method deduces them to a sequence one-dimensional (1D) ones by

$$\begin{aligned} &{}^{(i)}p^1(\mathrm{a}) = {}^{(i)}P_{\mathrm{p}}(x = \mathrm{a}),\\ &{}^{(i)}p^2_{\mathrm{a}}(\mathrm{b}) = {}^{(i)}P_{\mathrm{p}}(y = \mathrm{b} \mid x = \mathrm{a}),\\ &{}^{(i)}p^3_{\mathrm{a,b}}(\mathrm{c}) = {}^{(i)}P_{\mathrm{p}}(z = c \mid y = \mathrm{b} \ \&\ x = \mathrm{a}), \end{aligned} \tag{2}$$

where $i = 1, 2$. And then 1D histogram match method[14] can be applied on these 1D histograms which is formulated as

$$x \rightarrow T(x) = \min_t({}^{(1)}F(t) \geq {}^{(2)}F(x)), \tag{3}$$

where T is the transformation function needed to be calculated out, ${}^{(1)}F(t)$ and ${}^{(2)}F(t)$ are the cumulative histograms of the deduced 1D histograms of source image and style image, respectively.

In our approach, the foreground extracted object is treated as the target image and the background as the style image. Then the color harmonization problem can be solved via Neumann's method. HSV color space is used here to adapt the perceptual attributes. Furthermore, transfer on luminance or chroma channel is optional in HSV color space. The most common cases are to change the luminance of the foreground object with respect to the background. Under this condition, it is enough to perform hue-preservation color transfer, which only transfers the luminance channel while preserving the original hue channel. In a more complicated case, the background differs with the foreground object not only in luminance, but also in their chroma pattern. To such kinds of compositing problems, the complete color transfer on all three channels is needed.

HSV color space is a 3D cylindrical coordinate system, of which the hue channel is a periodic angle function whose period is 360°. Therefore, only a slight variation exists between hue values at 0° and 359°. If we transfer the hue channel by using Eq. (3) in the same way as S and V channels, colors at 0° and 359° would probably be mapped to totally different ones, thus losing its continuity. We solve this problem by introducing circular histogram-thresholding technique[15], which can find a optimal threshold repeatedly to divide the circular histogram in a linear histogram. After doing the circular histogram-thresholding on the hue channel, the histogram match can be applied according to Neumann's method.

Several groups of images, including natural ones and artistic works, are collected to evaluate the performance of the proposed image composition approach. When compositing natural images, the luminance channel transfer is usually enough. If there is a kind of color style in the background image, for example artistic

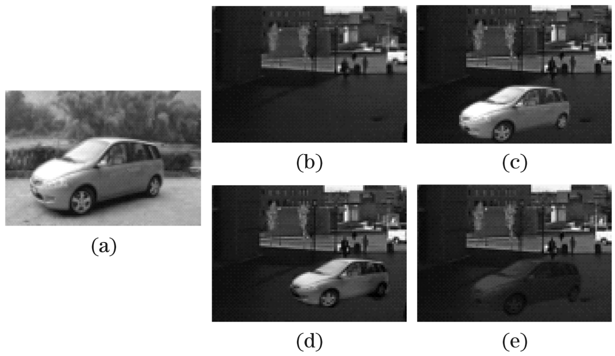

Fig. 1. Results of car and shadow, the simplest case of natural image composition. (a) Car, the foreground; (b) shadow of a building, the background; (c) result of direct paste; (d) result of Poisson editing; (e) result of the proposed method.

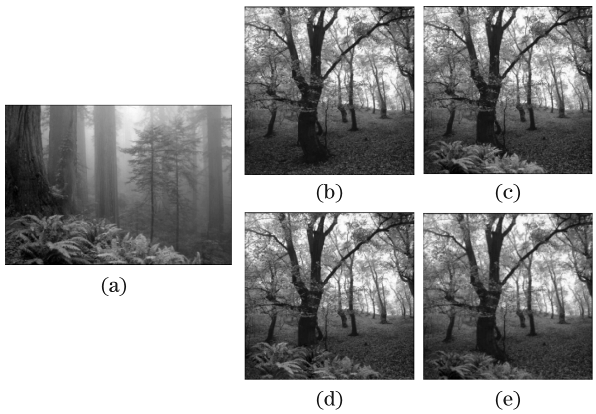

Fig. 2. Results of plants and forest. (a) Plant, the foreground; (b) forest, the background; (c) result of direct paste; (d) result of Poisson editing; (e) result of proposed method.

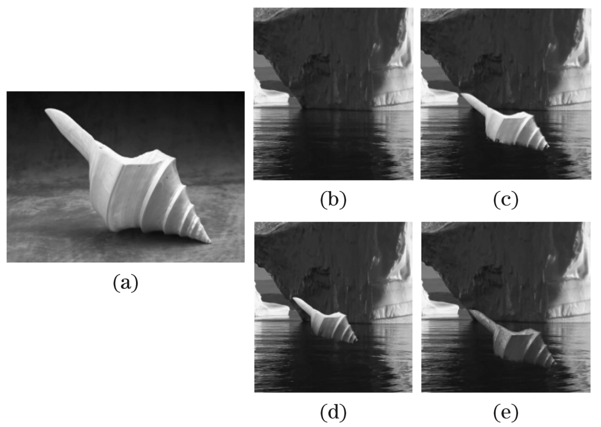

Fig. 3. Results of conch and sea, the art work example. (a) Conch, the foreground; (b) sea, the background; (c) result of direct paste; (d) result of Poisson editing; (e) result of proposed method.

works, however, the hue and saturation as well as luminance must be transferred. Thus the strict-optional 3D histogram match based color transfer is selected to fulfill the work of color harmonization.

Figures 1−3 show three result groups of image compositions considering color harmonization individually. Figure 1 shows a car in the sun to be composited in a shadow, which demonstrates the simplest case of hue-preservation composition. The result of direct paste exhibits obvious artifacts, as the car in the shadow of the building appears too bright. Poisson editing changes the luminance of the car at a certain degree, but the result is not satisfactory. Our method produces a reasonable luminance for this car in the shadow. In Fig. 2, the greenish plants are composited with the forest in autumn, and its color has to be changed to an "autumn color". The results of direct paste and Poisson editing are not convincing since the plants should be yellow in autumn. Our method extracts the plant from the foreground and then transfers the color style of forest in autumn to the plants. In Fig. 3, the big conch is composited with a sea image. In order to preserve the bluish style of the background (sea), and to preserve the affect of the shadow of the iceberg, both the color and the luminance of the conch should be transferred.

In conclusion, a novel digital image composition approach with color harmonization is proposed. This method consists of two parts: image matting and color transfer. Image matting is used to extract foreground objects, while color transfer is applied to make the foreground objects and the background share the same color characteristic. The proposed method is somewhat restricted by the abilities of matting and color transfer techniques. The matting methods nowadays cannot extract an object exactly from a complex natural image. Today, the color transfer approaches would lead to unacceptable results sometimes.

In the future, we hope to develop better matting and color transfer algorithms to produce more visually pleasing results even under worse conditions. We would also like to introduce the concept of image quality to this image composition system. Sometimes the differences between the foreground object and the background do not just exist in color style, but also in blur, contrast, texture component, and so on. This work will be extended to these aspects.

References

1. E. N. Mortensen and W. A. Barrett, in *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques* 191 (1995).
2. P. Guo and H. Lu, Acta Opt. Sin. (in Chinese) **22,** 1479 (2002).
3. C. Rother, V. Kolmogorov, and A. Blake, ACM Transactions on Graphics **23,** 309 (2004).
4. J. Jia, J. Sun, C.-K. Tang, and H.-Y. Shum, ACM Transactions on Graphics **25,** 631 (2006).
5. Y. Chen and C. Han, Chin. Opt. Lett. **5,** 25 (2007).
6. P. Pérez, M. Gangnet, and A. Blake, ACM Transactions on Graphics **22,** 313 (2003).
7. Y.-Y. Chuang, B. Curless, D. H. Salesin, and R. Szeliski, in *Proceedings of IEEE CVPR* 264 (2001).
8. J. Wang and M. F. Cohen, in *Proceedings of IEEE CVPR* 1 (2007).
9. Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum, ACM Transactions on Graphics **23,** 303 (2004).
10. L. Neumann and A. Neumann, in *Proceedings of Computational Aesthetics in Graphics, Visualization and Imaging* 111 (2005).
11. F. Pitié, A. C. Kokaram, and R. Dahyot, in *Proceedings of the Tenth IEEE International conference on Computer Vision* 1434 (2005).
12. S. Sun, Z. Jing, G. Liu, and Z. Li, Chin. Opt. Lett. **3,** 448 (2005).
13. T.-P. Wu, C.-K. Tang, M. S. Brown, and H.-Y. Shum, ACM Transactions on Graphic **26,** Article 8 (2007).
14. R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (2nd edn.) (Prentice Hall, Englewood Cliffs, 2002).
15. D.-C. Tseng, Y.-F. Li, and C.-T. Tung, in *Proceedings of ICDAR* 673 (1995).