Personal Photo Enhancement via Saliency Driven Color Transfer

Yuqi Gao, Jingfan Guo, Tongwei Ren, Jia Bei
State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
Software Institute, Nanjing University
{gyq1996, guojf}@smail.nju.edu.cn, {rentw, beijia}@nju.edu.cn

ABSTRACT

Personal photos on tour are easily affected by distractive objects, which requires effective post-processing for subject enhancement. In this paper, we propose a novel personal photo enhancement method using saliency driven color transfer, which can effectively reduce the attraction of distractive objects with simple user interaction. To each given image, distractive objects are firstly detected by combining saliency map and user interaction, and their attraction is further reduced by color transfer. The experimental results show that our method achieves similar effectiveness but higher efficiency to manually editing and outperforms other existing techniques.

Categories and Subject Descriptors

I.4.8 [**Scene Analysis**]: Color; I.4.9 [**Image Processing and Computer Vision**]: Applications

Keywords

Saliency detection; super-pixel representation; color transfer; photo enhancement

1. INTRODUCTION

Tourists prefer to take photos for memorizing tour experience and sharing them with friends [6, 14, 19]. Yet it is not easy to take a satisfactory photo in the places of historic interest and scenic beauty, for these places are usually crowded. Other tourists with colorful clothes in background may attract much attention of photo viewers and make the subjects less outstanding [15], who are called "*distractive objects*". Figure 1 shows some examples of distractive objects, which are marked with green boxes. These distractive objects are difficult to be completely avoided even the photographers notice them in time and retake photos from different views. Hence, some post-processing of personal photos is required to reduce their attraction for enhancing photo subjects.

ICIMCS '16, August 19-21, 2016, Xi'an, Shaanxi, China
© 2016 ACM. ISBN 978-1-4503-3528-7/15/08...$15.00
DOI: http://dx.doi.org/***********

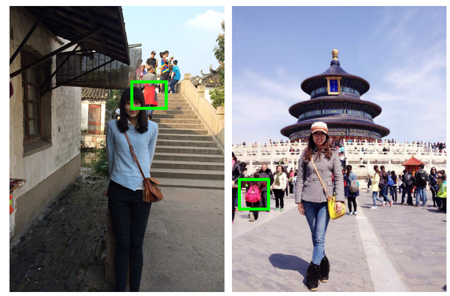


Figure 1: Examples of distractive objects (marked with green boxes).

Manually removing the distractive objects or transferring their color with software tools, such as Photoshop, is a common strategy to solve this problem. But it requires professional skills and it is usually time consuming. Recently, some semi-automatic techniques are proposed to reduce human labor in handling distractive objects. One kind of techniques focus on removing the distractive objects from the photos. Image inpainting is a typical technique for object removal, which aims to automatically fill the region of a removed object with the similar image content surrounding it or from other images [4, 17]. Image retargeting is another technique for object removal by insignificantly removing or warping the image content covering the removed objects [3, 13]. Both image inpainting and image retargeting can completely remove the distractive objects, but they easily bring in artifacts, especially for the personal photos with complex background. Another kind of techniques focus on changing the color of distractive objects to reduce their attraction. Color transfer aims to transfer the global or local color of a given image by referring some reference images or following pre-defined operations [8, 12]. It avoids to bring in artifacts in filling image content or changing image structure, but it requires suitable reference images or well-defined operations.

In this paper, we propose a novel personal photo enhancement method to reduce the attraction of distractive objects by saliency driven color transfer. Figure 2 shows an overview of our method. Given an input image with distractive objects, we first represent the photo with super-pixels, and calculate the saliency of each super-pixel. Meanwhile, we annotate the subject(s) in the photo with simple user interaction. Combining the saliency map and the annotated subject(s), the distractive objects are detected and their

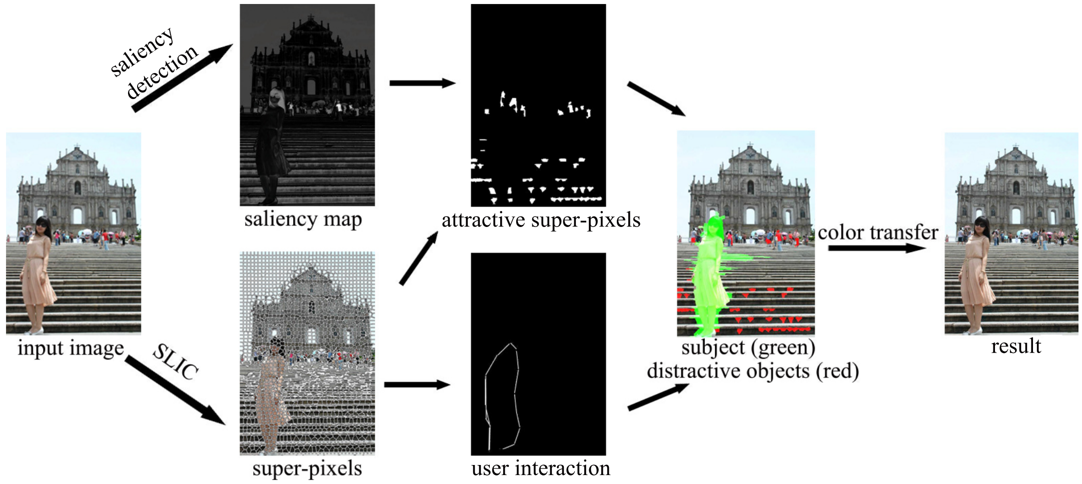


Figure 2: An Overview of the proposed method.

colors are further transferred according to the average colors of their surrounding super-pixels. In this way, the attraction of distractive objects are reduced and the subjects are enhanced. Though there has been existing work on combining saliency detection with color transfer [16] or image enhancement [10], to the best of our knowledge, our method is the first time to apply the combination of saliency and color transfer in handling distractive objects.

2. SALIENCY DRIVEN COLOR TRANSFER FOR DISTRACTIVE OBJECTS

2.1 Super-pixel Representation

Representing an image with super-pixels has shown its effectiveness in image editing [18]. In our method, super-pixels representation helps to annotate the subject(s) with user interaction, and it is also used as the preprocessing of distractive object detection and color transfer.

We utilize SLIC algorithm [2] for super-pixel representation, which can effectively generate a desired number of regular and compact super-pixels. It clusters pixels to super-pixels based on their similarity in $L^*a^*b^*$ space and spatial proximity. The distance measure d_s is calculated as follows:

$$d_s = \|\boldsymbol{\omega}_i, \boldsymbol{\omega}_j\|_2 + \frac{m}{\sqrt{N/K}}\|\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\|_2, \tag{1}$$

where $\|\boldsymbol{\omega}_i, \boldsymbol{\omega}_j\|_2$ denotes color distance in $L^*a^*b^*$ space and $\|\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\|_2$ denotes plane distance, here $\|\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\|_2$ is normalized by $\frac{m}{\sqrt{N/K}}$, in which m is a parameter to adjust the weight of spatial proximity, N is the number of pixels and K is the expected number of equally-sized super-pixels.

SLIC searches the optimal matching pixels from a square neighborhood around each initial seed, and generates super-pixel representation after iterations.

2.2 Saliency Detection

Salient object detection aims to detect the attractive objects for human in images [5, 7, 9, 11]. In our method, saliency map is utilized to detect the potential distractive objects. Specifically, the distractive objects are the salient objects that the users wish to avoid. Moreover, the color transfer is also conducted under the guidance of the saliency map.

We utilize a simple and effective salient object detection method called FT [1] in our method, which calculates the saliency value of each pixel $s_{x,y}$ as the Euclidean distance between its color and the average color of the Gaussian blurred image in $L^*a^*b^*$ space:

$$s_{x,y} = \|\boldsymbol{I}_{x,y} - \boldsymbol{I}_{avg}\|_2, \tag{2}$$

where $\boldsymbol{I}_{x,y}$ is the color of pixel $p_{x,y}$; $\boldsymbol{I}_{avg}$ is the average color of the Gaussian blurred image.

To each super-pixel sp_i, we calculate its global saliency as the average saliency value of all the pixels within it:

$$s_i^g = \frac{1}{N_i}\sum_{p_{x,y}\in sp_i} s_{x,y}, \tag{3}$$

where N_i is the number of pixels within sp_i.

We also detect the local saliency of each super-pixel by comparing its surrounding super-pixels:

$$s_i^l = \sqrt{\frac{1}{|\boldsymbol{\Delta}_i|}\sum_{sp_j\in\boldsymbol{\Delta}_i}(s_i - s_j)^2}, \tag{4}$$

where $\boldsymbol{\Delta}_i$ is the set of all the super-pixels surrounding sp_i, and $|\boldsymbol{\Delta}_i|$ is the number of super-pixels within $\boldsymbol{\Delta}_i$.

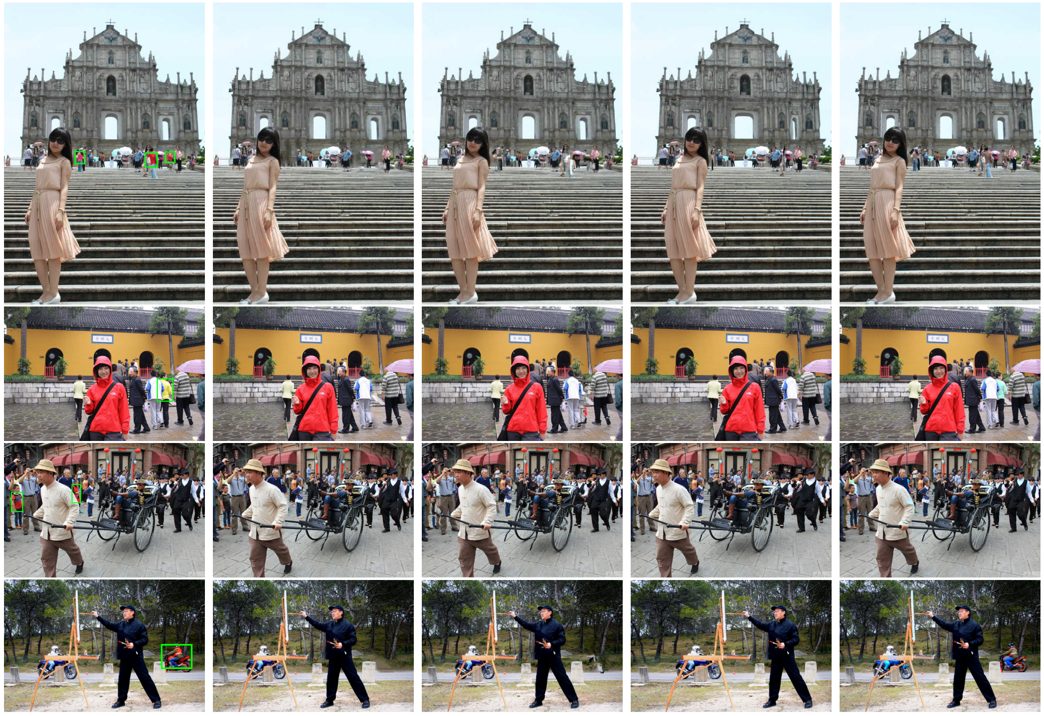
2.3 Salient-Driven Color Transfer

To detect the distractive objects, we first ask the users to manually annotate the subject(s) by drawing a curve around each subject. All the super-pixels inside the curve are treated as the subject.

Then, we select the salient super-pixels from the given image. Inspired by [1], we define a adaptive threshold T:

$$T = \frac{1}{W\cdot H}\sum_{x=1}^{W}\sum_{y=1}^{H} s_{x,y}, \tag{5}$$

where W and H are the width and height of the image. Based on T, we further define two thresholds $T^g = 2T$ and $T^l = T$ to measure global saliency and local saliency of super-pixels, respectively.



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

Figure 3: Comparison with other photo editing methods. (a) Input images. (b) Manual editing with Photoshop (PS). (c) Image inpainting (IP) [4]. (d) Seam carving (SC) [3]. (e) Our results.

All the super-pixels satisfy the following requirement are selected as the parts of distractive objects:

$$(s_i^g > T^g) \cap (s_i^l > T^l) \cap (sp_i \notin \mathbf{\Omega}), \tag{6}$$

where $\mathbf{\Omega}$ is the set of super-pixels annotated as parts of the subjects.

To each super-pixel selected as a part of distractive object, we reduce its attraction by saliency-driven color transfer. As shown in Equation (4) and (6), the selected super-pixel has obvious difference to its surrounding super-pixels. So we reduce its attraction by transferring its color to its surrounding super-pixels in $L^*a^*b^*$ space. In color transfer, we keep the color value on L channel of each pixel to avoid artifacts and change the color values on a and b channels of each pixel as follows:

$$a_{x,y}^{'} = (1 - \omega_i) \cdot a_{x,y} + \omega_i \cdot \tilde{a}_i, \tag{7}$$

$$b_{x,y}^{'} = (1 - \omega_i) \cdot b_{x,y} + \omega_i \cdot \tilde{b}_i, \tag{8}$$

where $a_{x,y}$ and $b_{x,y}$ are the color values on a and b channels of pixel $p_{x,y}$; $\tilde{a}_i$ and $\tilde{b}_i$ are the average color values on a and b channels of all the super-pixels surrounding sp_i, here pixel $p_{x,y}$ belongs to sp_i; ω_i is a weight parameter, which equals the local saliency s_i^l of sp_i in our method.

3. EXPERIMENTS

3.1 Experiment Settings

We first construct an image dataset consisting of 90 images with obvious distractive objects. These images are collected from the Internet and manually selected according to the existence of distractive objects. Based on the dataset, we compare our method with other photo editing approaches, including manual editing with Photoshop (PS), image inpainting (IP) [4], and seam carving (SC) [3].

The proposed method is implemented in Matlab. All of the experiments are carried out on a desktop computer with 2.7GHz CPU and 8GB memory.

3.2 Experimental Results

Figure 3 shows some examples of the results in our experiment, among which the distractive objects are marked with green boxes in the input images. For instance, in the first row of Figure 3, several people in red are detected as distractive objects. IP and SC introduce either white holes or serious distortion in removing the distractive objects, while PS obtains an acceptable result under professional operation but still causes blurring in the operation regions. In contrast, our method brings in less artifacts by changing the color of their clothes to gray, which effectively reduces the attraction of these distractive objects.

To quantitatively validate the performance of our method, twenty undergraduate students are invited to participate in user study. Each result generated by different methods on all the images is randomly evaluated by three participants at three levels: good, acceptable and bad. The dominant evaluation of three participants for each result is treated as the final evaluation, and a result is evaluated as acceptable if all the three participants have completely different opinions. Table 1 shows the result of user study. We can find that our method achieves similar effectiveness to PS and obviously outperforms the other two approaches.

We also evaluate the efficiency with the average running

Table 1: **Comparison of user study.**

	PS	IP	SC	Our
good	33	17	15	30
acceptable	56	45	52	53
bad	1	28	23	7

Table 2: Comparison of average running time.

	PS	IP	SC	Our
Language	–	Matlab	C++	Matlab
Time	25min	44.53s	2.57s	2.09s

time of each method, including user interaction. The evaluation result is shown in Table 2. We can find that our method is obviously faster than PS and IP and slightly faster than SC.

3.3 Discussion

Accurate user interaction is not required by our method. The users only need to roughly annotate the subjects, which is sufficient to locate the distractive objects. It is benefited from the super-pixel representation.

However, the performance of our method depends on the result of saliency detection. Once saliency detection fails, the distractive objects may not be detected. As shown in Figure 4, the blue umbrella marked by green box attracts much attention, but it is not detected in saliency map. It still exists in the generated result. This problem can be solved by introducing more user interaction.

4. CONCLUSION

In this paper, we propose a novel photo enhancement method to help users conveniently enhance their personal photos by reducing the attraction of distractive objects than directly removing them. The whole enhancement procedure is based on super-pixel representation and guided by saliency detection. The experimental results show that the proposed method can bring in less artifacts and achieve high efficiency.

5. ACKNOWLEDGMENTS

This work is supported by National Science Foundation of China (61321491, 61202320), Research Project of Excellent State Key Laboratory (61223003), Research Fund of the State Key Laboratory for Novel Software Technology at Nanjing University (ZZKT2016B09), and Collaborative Innovation Center of Novel Software Technology and Industrialization.

6. REFERENCES

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *CVPR*, pages 1597–1604. IEEE, 2009.

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *TPAMI*, 34(11):2274–2282, 2012.

(a) (b) (c)

Figure 4: An example of failure case. (a) Input image. (b) Binarized saliency map. (c) Our result.

[3] S. Avidan and A. Shamir. Seam carving for content-aware image resizing. 26(3):10, 2007.

[4] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *SIGGRAPH*, pages 417–424. ACM, 2000.

[5] A. Borji, M.-M. Cheng, H. Jiang, and J. Li. Salient object detection: A benchmark. *TIP*, 24(12):5706–5722, 2015.

[6] Y. Gao, J. Tang, R. Hong, Q. Dai, T.-S. Chua, and R. Jain. W2go: a travel guidance system by automatic landmark ranking. In *MM*, pages 123–132. ACM, 2010.

[7] J. Guo, T. Ren, and J. Bei. Salient object detection for rgb-d image via saliency evolution. In *ICME*. IEEE, 2016.

[8] Y. Hwang, J.-Y. Lee, I. S. Kweon, and S. J. Kim. Color transfer using probabilistic moving least squares. In *CVPR*, pages 3342–3349. IEEE, 2014.

[9] R. Ju, Y. Liu, T. Ren, L. Ge, and G. Wu. Depth-aware salient object detection using anisotropic center-surround difference. *SPIC*, 38:115–126, 2015.

[10] W.-M. Ke, C.-R. Chen, and C.-T. Chiu. Bita/swce: Image enhancement with bilateral tone adjustment and saliency weighted contrast enhancement. *TCSVT*, 21(3):360–364, 2011.

[11] Z. Liu, W. Zou, and O. Le Meur. Saliency tree: A novel saliency detection framework. *TIP*, 23(5):1937–1952, 2014.

[12] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley. Color transfer between images. *CGA*, (5):34–41, 2001.

[13] T. Ren, Y. Liu, and G. Wu. Image retargeting based on global energy optimization. In *ICME*, pages 406–409, 2009.

[14] J. Sang, C. Xu, and D. Lu. Learn to personalized image search from the photo sharing websites. *TMM*, 14(4):963–974, 2012.

[15] X. Sun, H. Yao, and R. Ji. What are we looking for: Towards statistical modeling of saccadic eye movements and visual saliency. In *CVPR*, pages 1552–1559. IEEE, 2012.

[16] J. Xia. Saliency-guided color transfer between images. In *ISVC*, pages 468–475. Springer, 2013.

[17] Z. Xu and J. Sun. Image inpainting by patch propagation using patch sparsity. *TIP*, 19(5):1153–1165, 2010.

[18] B. Zhang, R. Ju, T. Ren, and G. Wu. Say cheese: Personal photography layout recommendation using 3d aesthetics estimation. In *PCM*. Springer, 2016.

[19] Y.-T. Zheng, Z.-J. Zha, and T.-S. Chua. Mining travel patterns from geotagged photos. *TIST*, 3(3):56, 2012.