Novel Self-Portrait Enhancement via Multi-Photo Fusing

Sifeng Xia, Shuai Yang, Jiaying Liu* and Zongming Guo
Institute of Computer Science and Technology, Peking University, Beijing, China

***Abstract*—In this paper, we present a novel multi-photo-based framework to solve a self-portrait enhancement problem we call "1+2 problem", in which a self-portrait photo is enhanced with the help of two multiple photos that share the same scene and similar shooting time. The key idea is to exploit the extra information of these two photos to overcome the limited field of view and poor illumination of the target self-portrait. Given 1+2 photos, our method first performs an illumination adjustment to accomplish a uniform lighting style. Then these photos are aligned using the homography transformation matrices which are estimated by RANSAC over the matched SIFT feature points. Furthermore, a Markov Random Field based approach is proposed to globally fuse the aligned photos nicely. Experimental results demonstrate that the proposed method achieves satisfying results in this novel application scenario.**

I. Introduction

With the rapid development of digital life, sharing self-portraits on the social networks starts trending nowadays. However, these shared photos often fail to create the experience of re-visiting the scene, as self-portraits have particularly limited field of view (FOV). Indeed, when taking self-portraits, due to the short distance between the photographer and his camera, the photographer can take up as much as one-fourth to one-third area of a photo, making it hard to balance the foreground and background for a good photo composition. The background scene in self-portraits can be either incomplete or too small. Moveover, the poor illumination can also be a troublesome problem in self-portraits. Fortunately the photographer can easily take some other photos of the same scene after taking the self-portrait, and these photos (we call supporting photos) can help solve the limited FOV and illumination problems.

One way of solving the limited FOV problem is image inpainting technique. Image inpainting aims to reconstruct unknown region of an image. For FOV expansion, the outer region of a self-portrait is inpainted with the help of extra images. Data-driven-based image inpainting [1]–[3] retrieves reference images sharing similar scenes and combines them to fill the missing regions or extend the image boundary. In [1], similar images are searched based on gist features for inpainting, however the method assumes the matching scenes are roughly aligned but this assumption is usually not true in practice. Shan *et al.* [2] proposed a Markov Random Field (MRF) framework to align and fuse photos with the help of scene structure information obtained using structure from motion technique. And a more sophisticated graph-based image alignment method [3] is proposed to achieve better structure consistency. These methods carefully search matching images that share illumination consistency with each other from thousands of images. However, the self-portraits and supporting photos might have different lighting styles due to different shooting parameters between the normal camera and front-facing camera. The aforementioned methods may yield brightness variance artifacts in the results.

To unify lighting styles for visual consistency, illumination transfer methods have been studied. Traditional illumination transfer method is based on histogram matching, which yields artifacts in the result image if the source and target image differ a lot. Fang *et al.* [4] utilized edge preserving filter to transfer illumination of different parts of the source and target image separately. However, the results are distorted due to filtering process. Zhang *et al.* [5] casted color of the target image to the source image while transferring illumination, which is not suitable for images with different color distributions.

In this paper, we raise the "1+2" problem where one self-portrait photo is enhanced with the help of two supporting photos. To solve this problem, we propose a multi-photo-based framework combining lighting style unification and multi-photo fusion, which produces high-quality self-portraits that provide excellent experience of re-visiting the scene. We first perform lighting style unification by adjusting illumination of the photos with the guidance of one photo to accomplish global illumination consistency. By doing so, the brightness distinction in the final result is effectively avoided. Then the supporting photos are all aligned to the self-portrait using the homography transformation matrices estimated by RANSAC over the matched SIFT features. Finally the aligned photos are seamlessly fused under the proposed MRF energy framework. Experimental results demonstrate that our method can obtain satisfying results in this novel scenario.

The rest of this paper is organized as follows. Sec. II describes the proposed self-portrait enhancement method. Experimental results are shown in Sec. III and concluding remarks are given in Sec. IV.

II. Proposed Self-Portrait Enhancement Approach

Given a set of 1+2 photos $S_I = \{I_i\}, i = 0, 1, 2$ (one self-portrait photo I_0 and two supporting photos $\{I_i\}, i = 1, 2$)

* Corresponding author
This work was supported by National High-tech Technology R&D Program (863 Program) of China under Grant 2014AA015205 and National Natural Science Foundation of China under contract No. 61472011.

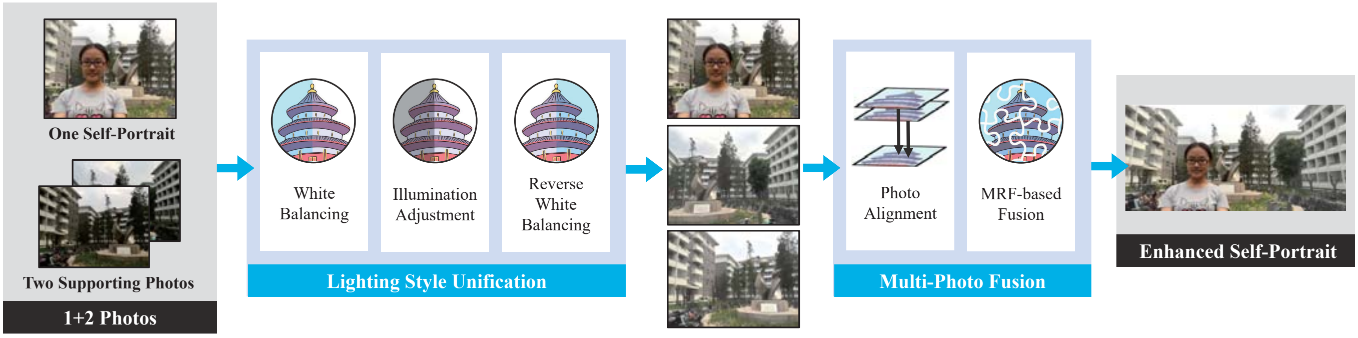


Fig. 1. Framework of the proposed multi-photo fusing method. Given one self-portrait and several supporting photos, our method first unifies their lighting style. A white balancing process is added to avoid color effects during the illumination adjustment and a reverse white balancing is performed to unify the color of the illumination. Then, the supporting photos are aligned to the self-portrait followed by a MRF-based fusion to nicely expand FOV of the self-portrait.

that share the same scene and similar shooting time, we first adjust their illumination to accomplish a uniform lighting style following the process of illuminant-aware color transfer approach [6]. Then we align supporting photos to the self-portrait I_0 using homography transformations. Finally the enhanced self-portrait $\hat{I}_0$ is obtained by fusing aligned photos together. $\hat{I}_0$ has a much broader FOV than I_0, resulting in better photo compositions. Fig. 1 shows the framework of our approach.

A. Lighting Style Unification

As is shown in the first row of Fig. 2, although the self-portrait I_0 and supporting photos are token in similar shooting time, there is still visible inconsistency in the lighting style. These inevitable inconsistency is mainly caused by the different shooting parameters between the normal camera and front-facing camera of a mobile phone. Besides, the photographer occupying large area of the scene in self-portraits can also affect the camera focus and exposure time, which aggravates the inconsistency. To solve this problem, before fusing, we first adjust the photos in S_I to make their lighting styles consistent with one chosen visually satisfying photo I_t in S_I.

During the illumination adjustment, we transfer the lighting style of the reference image I_t to other images I_s in S_I. Firstly, white balance adjustments are used to remove the color cast and bias in an image caused by different shooting parameters. Specifically, we compute the color adjustment parameters (r_w, g_w, b_w) for an image I using the weighted Grey-Edge algorithm [7], and divide the value (r, g, b) of each pixel with the color adjustment parameters:

$$(r, g, b) = (r/r_w, g/g_w, b/b_w). \tag{1}$$

After white balancing, we transform images I_s and I_t into YUV color space:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.148 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix}, \tag{2}$$

where Y channel is the illumination channel of the image and U, V are chrominance channels. Let Y_s and Y_t denote the illumination matrices (Y components) of I_s and I_t. Then we match the illumination of I_s to I_t using histogram matching with gradient preservations. The U and V channels remain unchanged.

In the first step of the illumination adjustment, histogram matching is used to convert illumination Y_s into intermediate illumination Y_m, which shares the same histogram as the reference illumination Y_t:

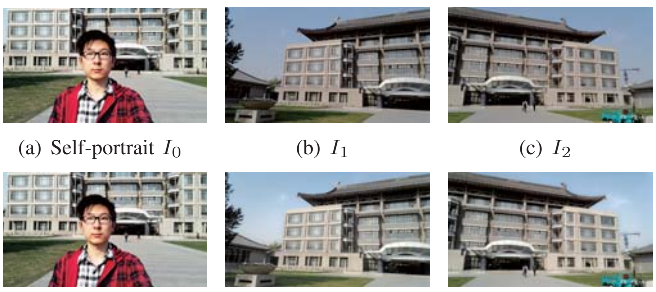
$$Y_m = \mathrm{C}_t^{-1}\left(\mathrm{C}_s\left(Y_s\right)\right), \tag{3}$$

where C_s and C_t are the cumulative histograms of Y_s and Y_t, respectively. $\mathrm{C}(\cdot)$ is the histogram equalization operation and $\mathrm{C}^{-1}(\cdot)$ is the corresponding reverse histogram equalization operation. In the second step, the output illumination Y'_s is obtained by taking the gradient preservation into account:

$$\left[\mathbf{I} + \lambda\left(\mathbf{D}_\mathbf{x}^\top \mathbf{D}_\mathbf{x} + \mathbf{D}_\mathbf{y}^\top \mathbf{D}_\mathbf{y}\right)\right] Y'_s = Y_m + \lambda\left(\mathbf{D}_\mathbf{x}^\top \mathbf{D}_\mathbf{x} + \mathbf{D}_\mathbf{y}^\top \mathbf{D}_\mathbf{y}\right) Y_s \tag{4}$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{D}_\mathbf{x}$ and $\mathbf{D}_\mathbf{y}$ are two gradient matrices along to x and y direction and λ is a weight to determine the relative importance of illumination transformation and gradient preservation.

Finally, a reverse white balancing is performed by multiplying the value of each pixel with the color adjustment parameters and the lighting-style unified I'_s is obtained. One example of lighting style unification is shown in Fig. 2. The lighting style inconsistency problem is effectively solved.



(a) Self-portrait I_0 (b) I_1 (c) I_2 (d) I'_0 (e) I'_1 (f) I'_2

Fig. 2. Lighting style unification results. Before lighting style unification, supporting photos (b) I_1 and (c) I_2 have poorer illumination than self-portrait (a) I_0. After lighting style unification, the illumination of all photos becomes consistent. In this example, I_0 is chosen as the target image I_t.

B. Photo Alignment

After obtaining lighting style unification photos $\{I_i'\}, i = 0, 1, 2$, we align each supporting photo I_i' to the self-portrait I_0'. Firstly, Scale-Invariant Feature Transform (SIFT [8]) feature points of I_i' are detected and their descriptors are computed. Then these feature points are matched between I_i' and I_0'. Next we perform a RANSAC-based voting algorithm over the matched feature points to find the best fitting transformation matrix $\mathbf{H}$. Finally we transform I_i' using the transformation matrix $\mathbf{H}$ and obtain I_i'' that is aligned to the self-portrait. After the transformation, all the supporting photos are aligned to I_0', as is shown in Fig. 3.

(a) Self-portrait I_0'

(b) I_1''

(c) I_2''

Fig. 3. Photo alignment results of the second row of Fig. 2. I_1'' and I_2'' are aligned to the self-portrait I_0'.

C. MRF-based Photo Fusion

In this section we fuse aligned supporting photos with I_0' into a single enhanced self-portrait $\hat{I}_0$. After alignment, photos have overlapped regions of the same scene and it should be carefully decided which photo to use in these regions of $\hat{I}_0$. Therefore, we formulate this problem as a MRF labelling problem, in which each photo represents a label and each pixel in $\hat{I}_0$ is assigned with a label. The value of pixel p in $\hat{I}_0$ with label i is set as the value of corresponding pixel p in I_i''. Our goal is to find the optimal labelling scheme.

Firstly, we initialize $\hat{I}_0$ as I_0', which ensures that the photographer in $\hat{I}_0$ is not covered by supporting photos. Let Ω denote the region where the pixel values are to be determined (for example, the black region in Fig. 3(a)) and $\delta\Omega$ be its boundary. Then given two labels, we define the MRF energy function to evaluate the labelling:

$$E(L) = \sum_{(p,q)\in\mathcal{N}_4} E_s(L(p), L(q)) + \alpha \sum_{p\in\Omega} E_d(L(p)), \quad (5)$$

where $L(p) = i$ means that the value of pixel p in $\hat{I}_0$ is decided by pixel p in I_i'', $\mathcal{N}_4$ is the neighborhood system and α is the weight to combine two energy terms:

Smoothness term: $E_s(L(p), L(q))$ penalizes the discontinuity within nearby pixels. It is defined as (for simplicity, we assume $L(p) = i$, $L(q) = j$):

$$E_s(L(p), L(q)) = ||I_i''(p) - I_j''(p)||_1 + ||I_i''(q) - I_j''(q)||_1, \quad (6)$$

where $I_i''(p)$ is the value of pixel p in I_i''.

Data term: $E_d(L(p))$ is defined as:

$$E_d(L(p)) = \begin{cases} +\infty, \text{if } I_i''(p) \text{ has no value} \\ 0, \text{if } I_i''(p) \text{ has value and } p \in \Omega\backslash\delta\Omega \\ ||P_i(p) - \hat{P}_0(p)||_1, \text{other} \end{cases}, \quad (7)$$

where $P_i(p)$ and $\hat{P}_0(p)$ are the patches centered at pixel p in I_i'' and $\hat{I}_0$ respectively. $||P_i(p) - \hat{P}_0(p)||_1$ is the patch difference measuring the consistency along the boundary of the self-portrait, and only known pixels in the patch are computed.

The energy optimization is achieved using multi-label graph-cuts algorithm. Furthermore, the Poisson fusion [9] is used to hide seams and to obtain the final $\hat{I}_0$.

III. Experimental Results

The proposed self-portrait enhancement method is implemented on Visual Studio 2013 platform. In the experiment, we set $\lambda = 1$ and $\alpha = 2$. To the best of our knowledge, there does not exist such a system that focuses on self-portrait enhancement and can achieve both unified illumination and expanded FOV. The closest one is the Photoshop PhotoMerge tool[1]. Thus in the experiment, we compare the performance of the proposed method with Photoshop. The self-portrait photos and the corresponding supporting photos for testing are taken by iPhone 6 Plus. Part of the testing photos are shown in Fig. 4. The whole photos and the experimental results have been released on our website[2].

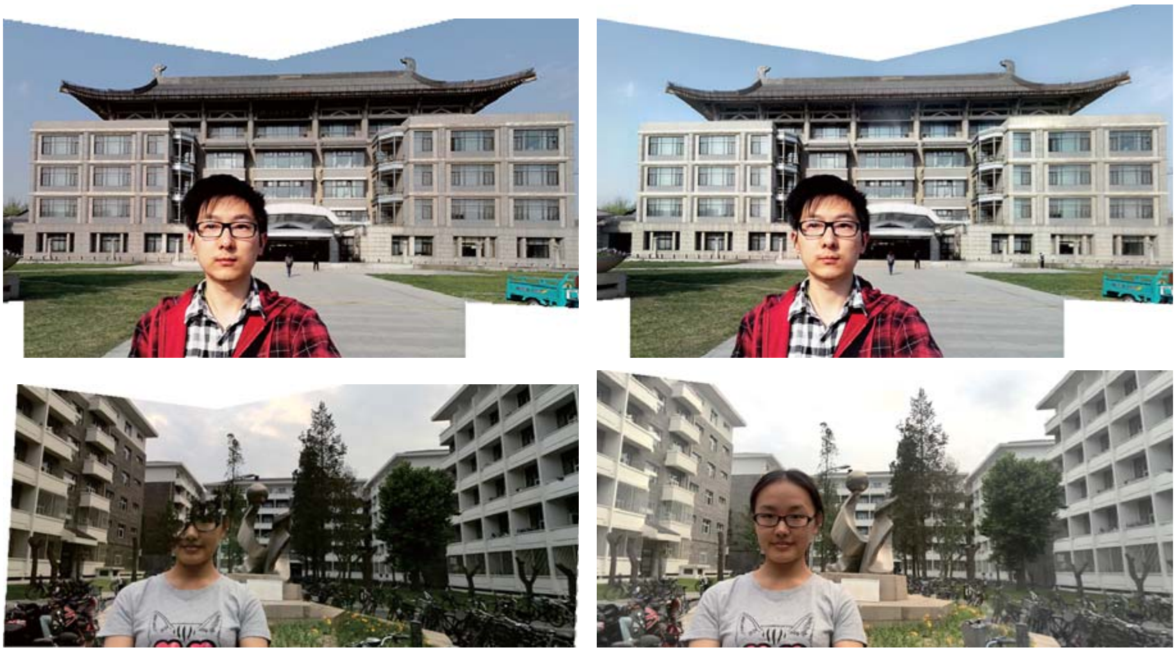
(a) Self-Portrait I_0 (b) Supporting photo I_1 (c) Supporting photo I_2

Fig. 4. Raw self-portrait I_0 and the corresponding supporting photos I_1 and I_2 for enhancement.

The self-portrait enhancement results are shown in Fig. 5. The changing areas of light and darkness are obvious in the Photoshop's result in the first row. This is because the input photos suffer great inconsistency in the lighting style and Photoshop only tried to smooth the local brightness changes, failing to achieve global illumination consistency. On the contrary, thanks to the lighting style unification, the proposed method obtained satisfying results with global illuminance consistency. In the second row of Fig. 5, although we have tried different settings of PhotoMerge tool, the girl in the Photoshop's result is always partly covered by the background scenes. By comparison, the proposed method gives self-portraits higher priority by initializing $\hat{I}_0$ as I_0', yielding plausible enhanced results.

[1]Adobe: Photoshop PhotoMerge, http://helpx.adobe.com/en/photoshop/using/create-panoramic-images-photomerge.html

[2]http://www.icst.pku.edu.cn/course/icb/Projects/SPEH.html



(a) Photoshop's results (b) Our results

Fig. 5. Comparisons with Photoshop PhotoMerge tool on self-portrait enhancement.

(a) Self-portrait I_0 (b) Supporting photo I_1 (c) Supporting photo I_2

(d) Enhanced self-portrait $\hat{I}_0$

Fig. 6. The performance of the proposed method on expanding FOV.

In Fig. 6, we show the performance of the proposed method on expanding a large FOV. The magnificent auditorium can be fully viewed in the enhanced self-portrait, creating satisfying experience of re-visiting the scene. In this example, the content-aware fill algorithm [10], [11] is used to fill the regions where no supporting photos can cover in $\hat{I}_0$.

IV. Conclusion

In this study, we raise the "1+2" self-portrait enhancement problem and propose a novel multi-photo-based framework to solve it. Given one self-portrait photo and two supporting photos that share the same scene and similar shooting time, we first adjust illumination of these photos to uniform their lighting styles. Then we utilize the homography transformations which are estimated by RANSAC over the matched SIFT feature points to align the photos. Furthermore, a MRF-based approach is proposed to fuse the aligned photos. Experimental results show that the proposed method can obtain high-quality self-portraits and create better experience of re-visiting the scene.

References

[1] J. Hays and A. A. Efros, "Scene completion using millions of photographs," *ACM Transactions on Graphics*, vol. 26, no. 3, July 2007.

[2] Q. Shan, B. Curless, Y. Furukawa, C. Hernandez, and S. M. Seitz, "Photo uncrop," in *Proc. European Conf. Computer Vision*, September 2014, pp. 16–31.

[3] Miao Wang, Yu Kun Lai, Yuan Liang, Ralph R. Martin, and Shi Min Hu, "Biggerpicture: data-driven image extrapolation using graph matching," *ACM Transactions on Graphics*, vol. 33, no. 6, pp. 1–13, 2014.

[4] Zhihong Fang, Hongyu Wu, Changjian Chi, and Xiaowu Chen, "Human head-shoulder illumination transfer," in *Proc. IEEE Int'l Conf. Virtual Reality and Visualization*, August 2014, pp. 122 – 129.

[5] Yi Zhang, Tianhao Zhao, Zhipeng Mo, and Wenbo Li, "A method of illumination effect transfer between images using color transfer and gradient fusion," in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, October 2013, pp. 1 – 6.

[6] Rang M. H. Nguyen, Seon Joo Kim, and Michael S Brown, "Illuminant aware gamut-based color transfer," *Computer Graphics Forum*, vol. 33, no. 7, pp. 319–328, October 2014.

[7] Arjan Gijsenij, Theo Gevers, and Joost van de Weijer, "Improving color constancy by photometric edge weighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 918–929, May 2012.

[8] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int'l Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, November 2004.

[9] P. Perez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Transactions on Graphics*, vol. 22, pp. 313–318, July 2003.

[10] Y. Wexler, E. Shechtman, and M. Irani, "Space-time completion of video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 463–476, March 2007.

[11] C. Barnes, E. Shechtman, A. Finkelstein, and D. B, "Patchmatch: A randomized correspondence algorithm for structural image editing," *ACM Transactions on Graphics*, vol. 28, no. 3, pp. 341–352, August 2009.