SALIENCE PRESERVING MULTI-FOCUS IMAGE FUSION

Richang Hong, Chao Wang, Yong Ge, Meng Wang, Xiuqing Wu, Rong Zhang

Dept of EEIS, University of Sci. & Tech. of China, Hefei, 230027, P.R.China
{richong, chaowang, ygesi, wangmeng}@mail.ustc.edu.cn {wuxq,zrong}@ustc.edu.cn

ABSTRACT

This paper proposes a novel multi-focus image fusion algorithm. Different from traditional decision map based methods, our algorithm is based on salience preserving gradient, which can better emphasize the structure details of sources while preserving the color consistency. We firstly measure the salience map of the gradient from each source, and then use their saliency to modulate their contributions in computing the global statistics. Gradients with high saliency are properly highlighted in the target gradient, and thereby salient features in the sources are well preserved. Furthermore we extend it to color domain by proposing an importance-weight based trigonometric average method to merge the color components. Extensive experiments on several datasets have demonstrated the effectiveness of our approach.

1. INTRODUCTION

Recently image fusion has become an important research topic in image analysis and computer vision. It is well known that *multi-focus* image fusion is an effective way to deal with the sensor's limitation of commercial cameras and extend the depth of defocus images. Thus it has drawn the attention of many researchers in the computer vision community. Currently most state-of-the-art multi-focus image fusion algorithms merge sources by combining the regions that are decided to be in better focus than their respective counterparts in the associated frames. But it is difficult to exactly judge whether the portion is in better focus. Furthermore multi-focus image fusion is expected to enhance the image quality. Thus algorithm that can extend the depth field of sources while emphasizing the structure details is highly desired to handle multi-focus image fusion.

In this paper, we propose salience preserving based algorithm which is different from the aforementioned algorithms. The algorithm is capable of emphasizing structure details, which is crucial in image enhancement. Moreover we have shown that the algorithm can be easily extended to color domain. The organization of the paper is as follows: In section 3, we discuss the construction of importance-weight based target gradient and briefly introduce the dynamic range compression. Section 4 presents a multi-focus color image fusion model, in which importance-weight based trigonometric average method is described in detail. Section 5 describes the experiments on a variety of multi-focus images we have performed and provides the results. And finally, we conclude in section 6, together with a discussion of future work.

2. RELATED WORK

Prior work of multi-focus image fusion mainly focused on gray scale images. A number of fusion strategies are proposed to solve the problem. Aizawa *et al.* [1] studied this problem and presented an algorithm based on camera point spread function (PSF). But it requires prior knowledge of the system, e.g. camera PSF. More algorithms are based on fusion decision map, i.e. merge sources by selecting portions which are in better focus, such as mosaic based [2][3] and multi-resolution based methods. The mosaic based methods select the portion in better focus in spatial domain directly. But it is sensitive to outliers. In multi-resolution, we usually have to apply wavelet [4][5] and discrete cosine transform, or pyramid based representation [6] to decompose the sources. Then from which select the sub-bands that satisfy the rules such as maximum energy and so on. But nonlinear operation on wavelet coefficients may induce the side effect of "ringing". Details can be seen around the edge of the book on the left in Fig. 1f.

3. SALIENCE PRESERVING FUSION

3.1. Gradient Fusion with Salience Map

We first adjust registered source images $\{f_k, k=1,2,\ldots,N\}$ to have the same mean gradients as the maximum mean gradient source for comparison and computation at the same level. For source f_k, we measure the salience of pixel p in f_k as follows [7]:

$$S_k^{'}(p) = mean_{q\in\theta_p}\{d(f_k(p), f_k(q))\}, \qquad (1)$$

$$S_k(p) = rescale\{1 - S_k^{'}(p)/\max_{q\in\Omega}(S_k^{'}(q))\}, \qquad (2)$$

where θ_p is the neighborhood of pixel p and Ω is the region of the whole image; rescale is an operation to ensure the dynamic range of S_k within [0,1]:

$$rescale(X) = (X - X_{\min})/(X_{\max} - X_{\min}),$$

and $d(a,b)$ is defined as

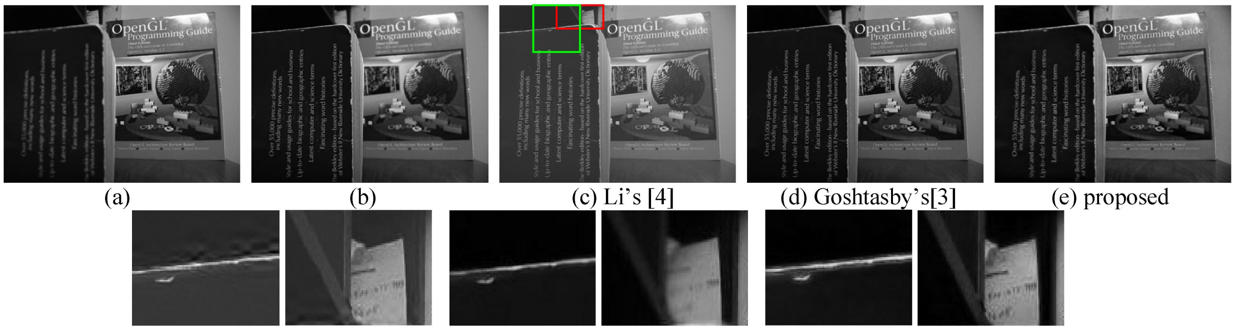
$$d(a,b) = e^{-(b-a)^2/2\sigma^2}.$$

In this paper, we set θ_p to be a 5×5 neighborhood of p, and $\sigma^2 = 100$. Such $S_k(p)$ represents the contrast around p and thus measures the local salience. We compare all the salience maps $S_k(p)$, and assign a normalized weight to each pixel in each source:

$$w_k(p) = \frac{S_k(p)^n}{\sqrt{\sum_{l=1}^{N}(S_l(p)^{2n})}}. \qquad (3)$$

Here, w_k is defined to be the importance weight of the source k. The positive parameter n reflects that on what degree the fused gradient resonates with the source of high salience. Then the importance-weighted contrast form is constructed as:

$$C(p) = \begin{bmatrix} \sum_k (w_k(p)\frac{\partial f_k}{\partial x})^2 & \sum_k w_k^2(p)\frac{\partial f_k}{\partial x}\cdot\frac{\partial f_k}{\partial y} \\ \sum_k w_k^2(p)\frac{\partial f_k}{\partial x}\cdot\frac{\partial f_k}{\partial y} & \sum_k (w_k(p)\frac{\partial f_k}{\partial y})^2 \end{bmatrix}. \qquad (4)$$



(a) (b) (c) Li's [4] (d) Goshtasby's[3] (e) proposed

(f) (g) (h) (i) (j) (k)

Fig. 1. Multi-focus gray scale image fusion: book. (a)(b) are sources, (c)-(e) are fusion results by Li's [4], Goshtasby's [3] and the proposed method, (f)-(k) are local images of the results respectively.

The target gradient ***V(p)***, at pixel *p*, is constructed by eigen-decomposition on *C(p)*. Such a target gradient ***V*** is actually the principle component of the sources' gradients weighted by their importance, which is the optimal representation for the weighted gradients in the sense of least-mean-square-error. But the fusion of gradient field may cause the dynamic range larger than the permitted one. So halo occurs when considering the dynamic range constraint in such case [8].To attenuate this effect, we adopt the following function to modify the target gradient, similar to [9].

$$V^*(p) = \left(\frac{\alpha}{|V(p)|}\right)^{1-\beta} \cdot V(p)\,. \tag{5}$$

Here β controls the strength of the boost for small gradient and α determines which gradient magnitudes remain unchanged (multiplied by a scale factor of 1). The parameter β is within (0,1) and set to 0.8, $\alpha = 0.8 \cdot mean\{|V|\}$ in the following section. Such a modified gradient ***V**** may result in halo reduction in general. Since the strong edges can also be easily observed, the target ***V**** preserves the salience, and we use it as the final target gradient of our proposed method.

3.2 Reconstruction from Target Gradient

Given the target gradient ***V*, t***he fused result can be found to be a 2D function *g* which minimizes:

$$\int_\Omega |\nabla g - V^*|^2 \, d\Omega, \quad g(x,y) \in [0,255]\,. \tag{6}$$

Such a function as (6) can be solved by iterative steps [10]:

$$\begin{cases} g(p)^{t+1/2} = g(p)^t + 1/4(\Delta g^t(p) - divV^*(p)) \\ g(p)^{t+1} = \max(0, \min(255, g(p)^{t+1/2})) \end{cases}. \tag{7}$$

Note that our approach is presented in a general mathematical form, so it can be implemented to an arbitrary number of sources. Furthermore, since it is based on the global statistics such as first-and-second order moment, it will be robust to outliers.

4. FUSION IN COLOR DOMAIN

Color image fusion should preserve color consistency [2]. In the case of fusing multi-focus color images, it is desired that the salience features, i.e. structure details, can be emphasized while its color consistency are preserved well.

4.1. Fusion Model

HSV color space has thee components, i.e., Value, Saturation, and Hue, which are relatively independent with each other. Hue represents the angle about the vertical axis of the cone, Saturation represents the ratio of the purity of Hue, and Value is the gray-scale of the color image. Among these three components, Value maintains the major structure details of a color image. Here we propose a fusion model for multi-focus color images based on *HSV* color space, as illustrated in Fig. 2.

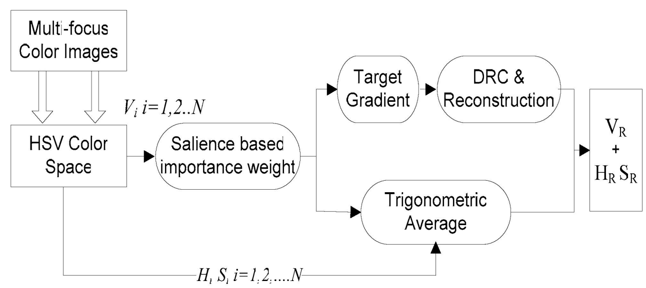


Fig. 2. Multi-focus color mage fusion model

As shown in Fig. 2, we denote the Value layers of input images as V_i, *i=1,2,…N*, they can also be regarded as the gray scale images extracted from color sources with different focus depth. The color layers of each source are denoted as H_i,S_i, *i=1,2,…N*. The first step is to measure the salience map of V_i, and then compute the importance weight based on saliency. After that, we use the weighted gradient to construct the contrast form, thus gradients with high saliency are properly highlighted in the target gradient. Dynamic range compression (DRC) is implemented on the target gradient to reduce halos. Finally we recover the composite Value component (i.e., gray scale image), which is denoted as V_R. More details about the algorithms mentioned above have been described in section 3. After fusing the Value components of sources with salience preserving method, the composite Hue and Saturation, denoted as H_R,S_R, are generated by importance-weight based trigonometric average algorithm, which is to be detailed in the following sub-section.

4.2. Importance-weight Based Trigonometric Average

In the context of multi-focus color images, since the sources are assumed to have the similar Saturation and Hue components, the average of Hue and Saturation from sources can be substitutes for the Hue and Saturation of the composite image. But when the Hue of sources close to Red (Hue=0) in different directions (as illustrated in Fig. 3), their average will be close to turquoise (Hue=180), i.e., induce color distortion.

Denote H_a and H_b as the Hue of the pixel in the same location of sources, and H_e as the expected composite Hue, as shown in Fig. 3. Let $H_f = (H_a+H_b)/2$. Then from the figure we can see that H_f is far away from H_e, which obviously not fits our goal. This thus introduces color inconsistency. An illustrative example is shown in Fig. 5e, where composite Hue is generated by amplitude averaging and color distortion can be clearly observed on the "book".

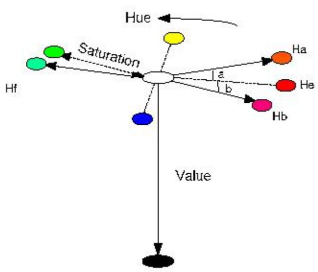


Fig. 3. Amplitude average

To deal with this problem, we propose an importance-weight based trigonometric average method which is based on two trigonometric functions. Denote the Hue component of the kth source image as H_k, then we let

$$H_R = \mathrm{Arg}\left(\sum_{k=1}^{N} w_k^2 \sin H_k, \sum_{k=1}^{N} w_k^2 \cos H_k\right), \qquad (8)$$

where w_k is defined in Eq. (3), Arg(., .) denotes the angle decided by the two components that are regarded as its sine and cosine values, and H_R is the composite Hue. Analogously we define S_k as the Saturation component of the kth source image, and the composite Saturation is derived by

$$S_R = \sum_{k=1}^{N} w_k^2 S_k . \qquad (9)$$

In this way the color distortion problem can be avoided. It is worth mentioning that the proposed approach is general and can be implemented with arbitrary number of source images.

5. EXPERIMENT RESULTS

We conduct various experiments to evaluate the proposed algorithm. Fig. 4 shows the result of the proposed algorithm applied on two gray scale images with different depth of field. We also illustrate the result obtained by Li *et al* [4] and Zhang *et al* [11] in Fig. 5c and Fig. 5d for comparison. From the figure we can see that our proposed algorithm outperform those in [4] and [11] in terms of preserving structure details (the results yielded by algorithms [4] and [11] have clear "ringing" effects).

Another example has been shown in Fig. 1. There the gray scale source images are extracted by keeping the Value components of the color images in Fig. 4. We can see that, similar to Fig. 4c and Fig. 4e, the image in Fig.1f also has the side effects of "ringing", which are induced by nonlinear operation on wavelet coefficients. On the other hand, this problem is greatly alleviated in our proposed and Goshtasby's algorithms (as illustrated in Fig. 1g and Fig. 1h respectively). However, Goshtasby's algorithm has another shortcoming in capturing structure details compared with our algorithm, as we can see that the "bookcase" on the right is blurry in Fig. 1i, whereas it is much clearer in our results.

In multi-focus color image fusion, the aforementioned "ringing" effects and shortcoming in capturing structure details exist as well in Goshtasby's and Bhagavathy's algorithms, which have been demonstrated by Fig. 5d and Fig. 5c. But we can see that our extended algorithm can successfully deal with these problems in color images as well (demonstrated by Fig. 5g). For comparison, we also illustrate the results obtained by performing amplitude average on Hue and Saturation and setting w_k to 0.5 simply in Fig. 5e and Fig. 5f respectively, where we can observe degraded color characteristics (there are many turquoise irregular shapes in Fig. 5e and the color on the top left corner of the right "book" varies from cyan to black in Fig. 5f). These results demonstrate the success of our proposed importance-weight based trigonometric average method. Fig. 6 illustrates another example, which further demonstrates the superiority of our proposed algorithm. As shown in the local images of the "balloon" in fusion results (Fig. 6c1 and Fig. 6c2), structure details (i.e., edges) and color consistency have been well preserved, whereas the result attained by Bhagavathy's algorithm has the side effects of "ringing" and color distortion.

6. CONCLUSION AND FUTURE WORK

This paper presents a salience preserving based multi-focus images fusion algorithm. We first measure the salience of sources and then construct the target gradient with dynamic range compression. Thus result image can be reconstructed from the target gradient. In this way, we demonstrate that salience features in the sources are well preserved. Moreover, we extend the algorithm to color domain and propose a multi-focus color image fusion model. Future works include further applying this framework to local image enhancement and extending the algorithm to fuse multiple frames of different exposure color images.

7. REFERENCES

[1] K.Aizawa, K.Kodama, and A.Kubota, "Producing object-based special effects by fusing multiple differently focused images", *IEEE Transactions on Circuits and Systems for Video Technology*, 2000. 10(2), pp. 323-330.

[2] Luca Bogoni, Michael Hansen, "Pattern-selective Color Image Fusion", *Pattern Recognition*, 2001, 34, 1515-1526.

[3] A.Goshtasby, "Image Fusion Systems Research" available at http://www.imgfsr.com/sitebuilder/images

[4] H.Li, B.S.Manjunath, and S.K.Mitra, "Multisensor Image Fusion Using the Wavelet Transform", *Graphical Models and Image Processing*, 1995, 57, (3), pp.235-245.

[5] G.Pajares, and J.M.Cruz, “A Wavelet-based Image Fusion Tutorial”, *Pattern Recognition*, 2004, 37, pp.923-931.

[6] P.R.Burt, and E.H.Adelsom, “The Laplacian Pyramid as a Compact Image Code”, *IEEE Trans. Comm*, 1983, 31(4), pp.540-583

[7] Y.F.Ma, H.J. Zhang, “Contrast-based Image Attention Analysis by Using Fuzzy Growing”, *ACM Multimedia*, 2003, pp. 374-381.

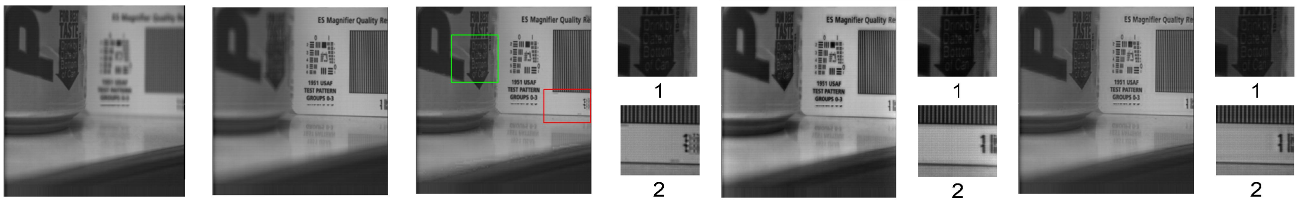
[8] Chao Wang, Qiong Yang, Xiaoou Tang, *et al* “Salience Preserving Image Fusion with Dynamic Range Compression”, *ICIP* 2006.

[9] R.Fattal, D.Lischinski, and M.Werman, “Gradient Domain High Dynamic Range Compression”, *ACM SIGGRAPH*, 2002.

[10] D.A.Socolinsky, “A Variational Approach to Image Fusion”, *Ph.D thesis*, *the Johns Hopkins University*, 2000.

[11] Z. Zhang and R. S. Blum, "A Categorization and Study of Multiscale-decomposition-based Image Fusion Schemes," *Proceedings of the IEEE*, Aug. 1999, pp. 1315-1328.

[12] Sitaram Bhagavathy, “Wavelet-based Multi-sensor Image Fusion”, available at http://vision.ece.ucsb.edu/~sitaram/ece594e.htm

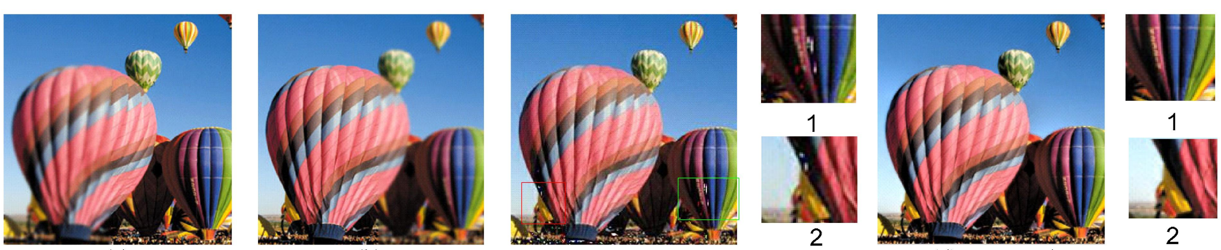


(a) (b) (c) Zhang’s (d) proposed (e) Li’s

Fig. 4. Multi-focus gray scale image fusion: Pepsi. (a)(b) are sources, (c)-(e) are fusion results by Zhang’s [11], the proposed and Li’s [4] method, (1) (2) are local images of the results respectively.

(e) (f) (g) proposed

Fig. 5. Multi-focus color image fusion: book. (a)(b) are sources, (c)(d)(g) are fusion results by Bhagavathy’s [12], Goshtasby’s [3] and the proposed method, (e)(f) are fusion results with color distortion.



(a) (b) (c) Bhagavathy’s (d) proposed

Fig. 6. Multi-focus color image fusion: balloon. (a)(b) are sources, (c)(d) are fusion results by Bhagavathy’s [12] and the proposed method, (1)(2) are images of the fusion results respectively.