An Image Harmonious Fusion Method Based on Two-Scale Decomposition

Yao Xu[1]

[1]School of Control and Computer Engineering, North China Electric Power University,
[2]Beinong Road, Huilongguan Street, Changping District, Beijing
578375123@qq.com

Abstract: ***A novel image fusion algorithm is introduced. In this method, the brightness channel of the image is decomposed first, and then the decomposed layers are processed in different ways. This method uses the mean interpolation method to solve the problem of image fusion. The experimental results show that the proposed method can obtain satisfactory results on images with different appearance, and the fusion results match the details and textures of the target image perfectly and seamlessly. In addition, the parallel implementation based on Graphics Processing Unit (GPU) improved the efficiency of the algorithm.***

Keywords: Image fusion, Image harmony, Mean coordinate, Image inpainting

1. Introduction

Image fusion is a useful operation in image and video editing applications that aims to produce realistic fusion results. In recent years, some related image fusion techniques have been proposed in the field of computer graphics and image, among which the image fusion technique based on gradient domain is a simple and widely used algorithm in image editing. Gradient domain technology solves the Poisson equation in the gradient domain to obtain the seamless fusion result, and solving the image fusion problem in the gradient domain takes advantage of the characteristics that people often pay more attention to the intensity difference of the local image than the intensity itself when observing the image. The seamless blending technique based on gradient domain was first proposed by Perez et al. [1], Jia et al. [2] improved the blending result by optimizing the blending boundary, and Lalonde et al. [3] introduced a blending mask to remove the boundary stain. Tao et al. [4] proposed a fault-tolerant method to prevent color change without changing the position of the boundary, and Chen et al. [5] proposed an image synthesis method combining improved Poisson fusion and alpha blending.

Recently, 2-D barycentric coordinates -- mean coordinates are used to solve the interpolation problem in image fusion. Farbman et al. [6] obtained very similar Poisson image fusion results using the mean coordinate method. Lee et al. [7] proposed an improved sampling method to solve the mean coordinate concave domain problem that Farbman et al. [6] did not solve, and some methods [8],[9] that combined mean coordinate and matting techniques were also proposed. In addition, Wang et al. [10] introduced the harmonic fusion method, which uses harmonic coordinates instead of mean coordinates in image fusion.

The above methods can usually obtain good results of image fusion. However, if the source and target images have different appearances (e.g., captured with different cameras under different conditions), the techniques based on the gradient domain and the methods based on coordinates will not be able to obtain realistic results. Obviously, the ideal result satisfies the following goals: it should have no edges and match the target image visually. Recently, Sunkavalli et al. [11] explicitly solved the problem of image harmonization in image fusion. They extend the gradient-domain technique by reconstructing the image in one large filter output and incorporating harmonization into the resultant result. Their method uses multiscale techniques to automatically correct inconsistencies in the image.

Harmonious image fusion requires that the global appearance of the synthesized region is harmonious with the background image. This involves three aspects: first, the brightness of the result image should match that of the target image. For the regions with similar structure, the brightness of the result image and the target image should be similar in these regions; Secondly, the detail texture of the synthesized image should be realistic. Finally, the color of the resultant region should be harmonious. We observe that in most cases, if the source and target images have similar structures, details and colors in the fused region, the fused image looks more harmonious. However, if the structures in the fused regions are different and the colors of the source images are preserved, the fused results usually look more harmonious.

Based on the above considerations, a novel image fusion method is proposed in this paper. This method uses different ways to satisfy the above three aspects. Firstly, the algorithm decomposes the image into a base layer, a detail layer and a color layer. In order to obtain a harmonious intensity distribution on the synthesized image, the intensity statistical characteristics of the target image base are transferred to the base of the source image. The realistic details of the fused image are obtained by region inpainting technique and weighted blending technique. Once the harmonious luminance layers are calculated, the algorithm uses the mean fusion technique to generate the final composite result on each layer. Because the mean fusion of each pixel is completely independent of other pixels in the fusion region, the algorithm uses a parallel method to achieve mean fusion to speed up the whole process.

2. Algorithm Overview

The proposed algorithm is divided into four main steps. Firstly, the source image and target image were decomposed into three layers: base layer, detail layer and color layer. Secondly, different methods are used to process the information of each layer. Firstly, the brightness histogram of the target image is transferred to the base level of the source image to achieve brightness balance, and the hole completion method is used to modify the detail layer of the target image in the fusion region, and the resulting details are obtained. Third, mean fusion is used on each pair of corresponding layers; Finally, the final result is obtained by synthesizing the results of the three layers. The work flow on the brightness layer is shown in Fig. 1.

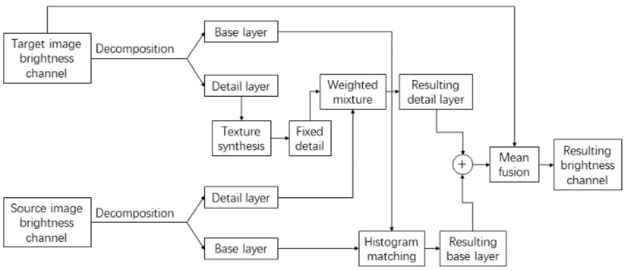

Figure 1 Workflow of brightness layer

2.1. Layer decomposition

The source and target images are first converted to CIELAB color space, and then the luminance channel is decomposed into base and detail layers. The base layer is the structure layer which contains the structure information of the image. Edge smoothing methods such as bilateral filtering [12] and weighted least squares (WLS) filtering [13] can be used to decompose the base and detail levels. The proposed algorithm uses the weighted least squares method because of its excellent performance in resolving blurring levels close to edges.

Given a luminance channel L, the base B in L can be obtained by solving the minimum value below:

$$\sum_p \big(B(p)-L(p)\big)^2 + \lambda\left(\frac{|B_x(p)|^2}{|L_x(p)|^{\alpha}+\varepsilon} + \frac{|B_y(p)|^2}{|L_y(p)|^{\alpha}+\varepsilon}\right) \qquad (1)$$

Where, p represents each pixel in the fusion area,B represents the smoothed base layer, L represents the image brightness channel $L_x(p), B_x(p)$ represents the gradient in the x direction, $L_y(p), B_y(p)$ represents the gradient in the y direction. The first term aims to keep B as similar as possible to L The second term smooths B by minimizing its partial derivative. The coefficient λ is used to balance these two terms: the larger the value of λ, the smoother the base layer will be. The subscripts x and y stand for taking partial derivatives on the abscissa and ordinate, respectively. The coefficient α determines the sensitivity of the L gradient. In order to prevent divisibility by 0, the coefficient ε is set to 0.0001. In the experiments in this paper, $\alpha = 1.2$ and $\lambda = 1$.

Subtraction is then used to obtain the detail layers of the source and target images, respectively.

$$D = L - B \qquad (2)$$

Where D represents the detail layer of the image.

2.2. Brightness matching

The base layer contains the distribution information of the brightness space. The algorithm tries to load the intensity spatial distribution of the target image into the source image so that it can get the ideal result when matching the fusion region. Histogram matching is a main solution to modify the intensity of the source image so that the intensity statistics of the source image can be similar to that of the target image. Given a source image base base Bs with histogram hbs and a source image base base Bh s with histogram hbt harmony, it can be calculated by the following formula:

$$B_s^h = histmatch(B_s, h_{bt}) \qquad (3)$$

Where B_s represents the base layer of the source image h_{bt} represents the histogram of the base layer of the icon image B_s^h represents the harmonious source image base layer. Here, $histmatch$ () translates the h_{bt} of the target image to the B_s of the source image.

Histogram matching is a powerful tool for image appearance matching [11]. Since the proposed algorithm uses histogram matching on the base layer with less noise and details, it will minimize some defects such as noise amplification.

2.3. Detail manipulation

If the detail layer of the source image and the detail layer of the target image are directly fused together, the fusion region will contain a lot of detail layer structure information of the target image, resulting in defects. The result of coordination should be that the fusion region does not contain the original structural details of the target image in the fusion result. An intuitive way is to separate the structural information from the detail layer of the target image. But directly separating this information in a separate layer is difficult compared to the previous transformation. Therefore, the algorithm assumes that the fusion result only needs to be visually realistic, which means that the pixels in the fusion region need to match the details of the surrounding pixels in appearance. Based on this assumption, the algorithm uses two steps to deal with detail reconciliation. Firstly, the detail part of the fused region is removed from the detail layer of the target image and the hole is filled. Then, the source image detail layer and the target image detail layer are mixed together in a weighted way to obtain the result detail layer.

Inpainting the holes in the target detail image is to fill those missing pixels. There are two ways to solve this problem: texture synthesis and image inpainting methods. Texture synthesis works well when dealing with random 2D patterns. Image inpainting, on the other hand, works on linear structures which can be thought of as one-dimensional patterns. In this paper, we adopt a method based on Criminisi et al. [14], which combines the advantages of texture synthesis and image inpainting. In order to produce reasonable results, the algorithm uses the image pixels around the edge of the hole as texture patches and obtains the modified target detail D_t' in the fusion region.

$$D_r = \gamma_s D_s + \gamma_t D_t' \quad (4)$$

Where D_r represents the detail layer of the result image D_s represents the detail layer of the source image and D_t' represents the detail layer of the target image after hole restoration. These two coefficients $\gamma_s, \gamma_t \in [0,1]$ control the proportion of each detail. Different values of γ_s and γ_t can produce different fusion effects. For seamless fusion because the result only includes the details of the source image and the details of the target image should be hidden. Then $\gamma_t = 0$ can be taken to hide D_t' and $\gamma_s = 1$ can be taken to convert D_s to D_r. In the application of harmony fusion in this paper, $\gamma_t > 0$ can be taken to preserve some details of the target image. Note that the sum of the weights of the two cannot be too small otherwise the result will look unreal due to the lack of too much detail.

2.4. Parallel mean fusion

Once the base B_s^h of the source image and the blended detail layer D_r Are obtained, the new source image brightness channel L_s' can be calculated by $L_s' = B_s^h + D_r$. Then the new brightness channel is fused to the brightness layer of the target image by the mean coordinate method.

Given a region Ω to be fused and its chain $\partial\Omega(t_0 t_1, \dots, t_m = t_0)$ of boundary pixels given in counterclockwise order, the mean coordinate (MVC) of each pixel $p \in \Omega$ is the weight k_i. k_i is calculated using the following formula:

$$k_i = \frac{\omega_i}{\sum_{j=0}^{m-1} \omega_j}, i = 0,1, \dots, m-1 \quad (5)$$

$$\omega_i = \frac{tan(\beta_{i-1}/2) + tan(\beta_i/2)}{||t_i - p||} \quad (6)$$

Whereβ_i is $\angle t_i p t_{i+1}$ Figure 2 shows the definition of β_i.

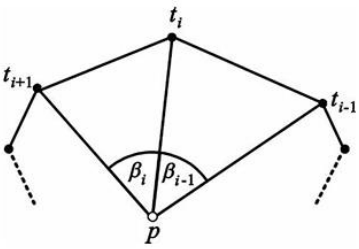

Figure 2 Illustration of mean value coordinate angle

The mean interpolation $r\ (p)$ using the mean coordinates of these pixels can be defined as follows:

$$r(p) = \sum_{i=0}^{m-1} k_i(p)\left(L_t(t_i) - L_s'(t_i)\right) \quad (7)$$

The resulting luminance layer L_r is then defined as

$$L_r = L_s' + r \quad (8)$$

Because the mean interpolation of each pixel in the fusion region is only related to the boundary coordinates, the interpolation of each pixel can be handled independently. To solve this performance bottleneck, we use CUDA technology to accelerate interpolation.

The previous results obtained with the formula $L_r = L_s' + r$ may be outside the brightness range [0, 100]. This may lead to a loss of visual content in the image so these values should be converted to the display range. The algorithm successfully preserved the rich details of the result image by migrating the intensity histogram of the target intensity L_s to the current intensity L_r.

2.5. Color layer fusion

In the color channel fusion algorithm, a_s and b_s of source images are directly used for mean fusion processing. Finally, the algorithm converts the image from CIELAB color space to RGB color space and obtains the final harmonious fusion result.

3. Experiments and Results

In this paper, a novel method of harmonious image fusion is implemented. The proposed method is implemented using MATLAB on a computer configured as: CPU (Inter Core2Duo 2.66GHz), 2G main memory, and graphics card with 1G graphics memory NVIDIAGeForce GTS250. This method requires two parameters: γs and γt. Different fusion objectives can be obtained with different parameters.

The most time-consuming steps in the implementation of the algorithm are region inpainting and calculating the Mean Value Coordinate (MVC) of each pixel. Because the mean interpolation for each pixel is computed independently of all other pixels, this step is ideal for parallel implementation on a Graphic Processing Unit GPU. The performance of the algorithm in calculating the mean coordinates is shown in Table 1 where N-b represents the number of boundary points of the fusion region and N-r represents the number of pixels inside the fusion region. Note that the GPU runtime in the table includes disk input and output.

Table 1 The mean value coordinate performance

experiment	*N_b*	*N_r*	*CPU/s*	*GPU/s*
footprint	989	64337	10.801	1.372
face	1007	85863	14.792	1.695

The proposed method is compared with the original mean fusion method. The footprints of the source image in Figure 3 are cropped out and fused into the target image Figure 3 (c) shows the result produced using the original mean fusion method. Although the result is seamless, the fusion result lacks similar details to the target image, which leads to a discordant appearance between the fused region and other pixels and the image looks unnatural. In the proposed method, weighted blending of region inpainting and detail is appropriately applied to produce seamless boundaries and obtain realistic detail effects so that the results obtained are harmonious and the results are shown in Figure. 3 (d).

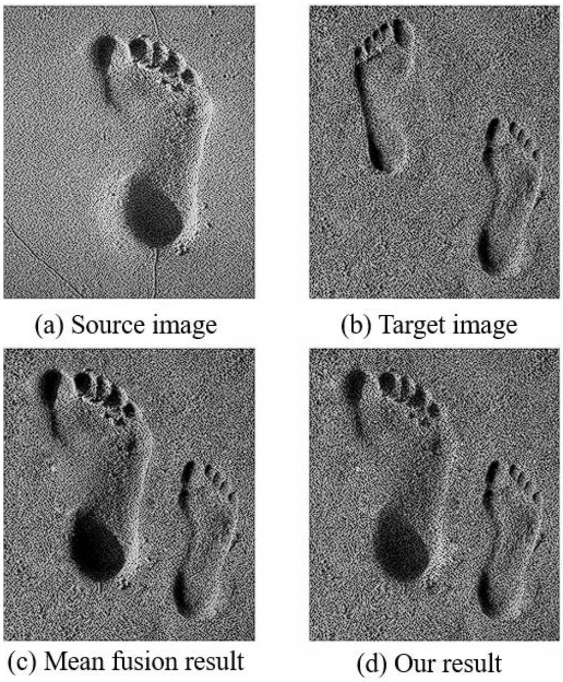

Figure 3 Footprint in source image

Although the original mean fusion works well in most cases, the proposed algorithm provides features that are lacking in the original mean fusion. Detail blending is crucial in achieving harmonious results. Figure 4 shows another example: the source image face is fused into the target image. The original mean fusion method produces less realistic results as shown in FIG. 4 (c). The results of this paper, shown in FIG. 4 (d), show visually realistic detailed textures in the fused region. As shown in Figure 4, the new method proposed in this paper can produce more realistic results than the original mean fusion method.

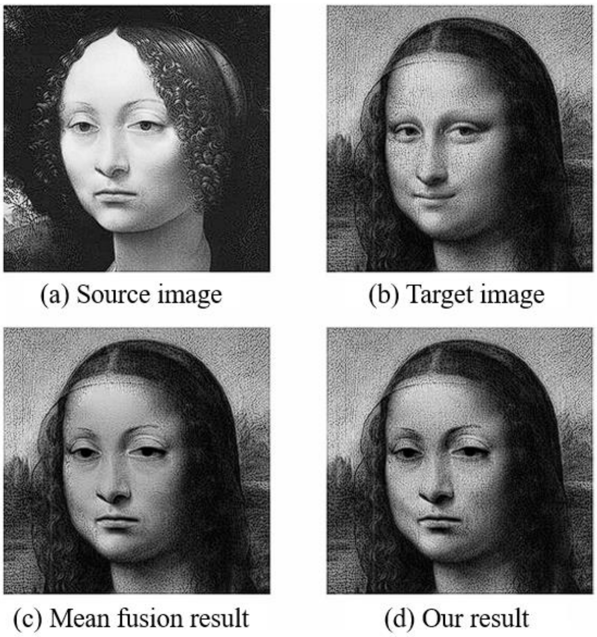

Figure 4 An example of harmonious cloning

Note that the result of the MVC algorithm is too smooth while the result of the proposed algorithm has more reasonable texture details.

4. Conclusion

This paper presents a novel algorithm for harmonious image fusion based on two-scale decomposition. The method uses region-filling and weighted blending methods to produce visually realistic fused image details. The experimental results show that the proposed method can match the target image well in detail.

Despite the above advantages of the algorithm, it still has some limitations such as the current implementation is only partially parallel and region filling consumes a lot of CPU time. Therefore, we should try to use GPU to process the whole process to improve the computational efficiency of the algorithm in the future work.

References

1] Pérez P Gangnet M Blake A . Poisson image editing [J] ACM Transactions on Graphics 200322 (3):313-318
2] Jia J Sun J Tang C-K et al . Drag-and-drop pasting [J].ACM Transactions on Graphics200625 (3):631-637
3] Lalonde J-F Hoiem D Efros A A et al . Photo clip art [J].ACM Transactions on Graphics 2007 26 (3):1-10 .
4] Tao M W Johnson M K Paris S . Error-tolerant image compositing [C] / / Proceeding of the 11th European Conference on Computer Vision . Heraklion Crete Greece: IEEE2010: 31-44 .
5] Chen T Cheng M M TanP et al . Sketch2photo: internet image montage [J].ACM Transactions on Graphics 2009 28 (5):124-134 .
6] Farbman Z Hoffer G Lipman Y et al., Coordinates for instant image cloning [J]. ACM Transactions on Graphics, 2009 28 (3):1-9
7] Lee S-Y Lee I-K . Improved coordinate-based image and video cloning algorithm [C] / / Proceedings of ACM SIGGRAPH Asia . New York: ACM 2009: 1-5 .
8] Xie Z-F ShenY MaL-Z et al . Seamlessvideo composition using optimized mean-value cloning [J]. The Visual Computer 2010 26 (6-8):1123-1134 .
9] Ding M Tong R-F . Content-aware copying and pasting in images [J]. The Visual Computer 2010 26 (6-8): 721-729 .
10] Wang R Chen W-F Pan M-H et al . Harmonic coordinates for real-time image cloning [J]. Journal of Zhejiang University Science 2010 (11): 690-698
11] Sunkavalli K Micah K J Matusik W et al . Multi-scale image harmonization [J].ACM Transactions on Graphics 201029 (4):125-134 .
12] Tomasi C, Manduchi R . Bilateral filtering for gray and color images [C] / / Proceedings of the Sixth International Conference on Computer Vision . Bombay India: IEEE 1998: 1-5 .
13] Farbman Z Fattal R Lischinski D et al . Edge-preserving decompositions for multi-scale tone and detail

manipulation ［J].ACM Transactions on Graphics 2008 27（3): 1-10 .

14] Criminisi A Pérez P Toyama K . Region filling and object removal by exemplar-based inpainting ［J]. IEEE Transactions on Image Processing 200413（9): 1200-1212