International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 31, No. 4 (2017) 1754006 (14 pages)

DOI: 10.1142/S0218001417540064

An Image Stitching Algorithm Based on Histogram Matching and SIFT Algorithm

Jing Zhang[*,†,§], Guangxue Chen[*,¶] and Zhaoyang Jia[*,‡]

[*]*State Key Laboratory of Pulp and Paper Engineering*
South China University of Technology, Guangzhou 510640, P. R. China

[†]*College of Arts, South China Agricultural University*
Guangzhou 510640, P. R. China

[‡]*College of Packaging & Printing Engineering, Tianjin*
University of Science & Technology, Tianjin 300222, P. R. China
[§]*smilejinger@126.com; zhangjing287@163.com*
[¶]*chengx@scut.edu.cn*

Received 28 May 2016
Accepted 1 August 2016
Published 24 October 2016

Image stitching among images that have significant illumination changes will lead to unnatural mosaic image. An image stitching algorithm based on histogram matching and scale-invariant feature transform (SIFT) algorithm is brought out to solve the problem in this paper. First, histogram matching is used for image adjustment, so that the images to be stitched are at the same level of illumination, then the paper adopts SIFT algorithm to extract the key points of the images and performs the rough matching process, followed by RANSAC algorithm for fine matches, and finally calculates the appropriate mathematical mapping model between two images and according to the mapping relationship, a simple weighted average algorithm is used for image blending. The experimental results show that the algorithm is effective.

Keywords: Image stitching; histogram matching; SIFT; feature matching; image blending.

1. Introduction

In order to obtain the ultra-wide-angle and high-resolution images, the traditional way is to use expensive special camera equipment to photograph and process images. But in recent years, with the popularity of digital cameras, smart phones and other economic hand-held imaging devices, it is convenient for people to get discrete image sequence for some scenes, and then through the appropriate image processing methods to improve the quality of the image and finally, realize automated image stitching. This approach can also create the ultra-wide-angle and high-resolution mosaic image.

[¶]Corresponding author.

Image stitching is the process of combining multiple photographic images to produce a segmented panorama or high-resolution image. If there are two images with overlapping area, the image stitching is a mosaic of these two images into one image.[9] The key of image stitching is the ability to quickly and efficiently find the overlap of two different images for image alignment, thus achieve a wide-angle image. Lowe published the scale-invariant feature transform (SIFT) algorithm in 1999, SIFT algorithm is an algorithm in computer vision to detect and describe local features in images, Lowe summed up and improved it in 2004.[7] Szeliski Richard proposed the new theory of image registration, perform image enhancement before image registration so that the feature points in the image become obvious and improve the accuracy of registration.[10]

In this paper, histogram matching and SIFT algorithm are used in image stitching, which can improve the quality of the mosaic image, especially for images that have significant illumination changes. Image stitching among images that have significant illumination changes will lead to visible transitions and the mosaic image looks unnatural, therefore this paper presents a preprocessing method — histogram matching first; then SIFT algorithm is performed for image alignment; the weighted average blending method is used to remove the splicing gap and ultimately achieve seamless photo-mosaics.

2. Theory of Image Stitching

The image stitching process can be divided into four main stages, which are image acquisition, image preprocessing, image alignment and image blending.

2.1. *Image acquisition*

Image acquisition is usually the first step of image processing system. It is the process to get the original image. There are many ways to acquire images, different ways are chosen according to different applications, for example, people use telescope to take beautiful photos of the universe, medical research obtained microbial images with a microscope, digital cameras or cell phones are used to take pictures of everyday life.

2.2. *Image preprocessing*

During the process of digital image acquisition by hardware, there are different interferences on the acquired original image, so the quality of the acquired image cannot meet the expected. The original image should be preprocessed effectively and accurately such as image denoising, image correcting and so on, to ensure a precise pixel-level calibration accuracy between images. The main purpose of image preprocessing is to reduce the difficulty of image alignment. The accuracy of image preprocessing stage has a direct impact on the quality of the final mosaic image.

2.3. *Image alignment*

The quality of mosaic image mainly depends on the accuracy of image alignment. Image alignment is the process of image processing among images with overlapping area, which may be acquired from different sensors, times, depth or viewpoints. It involves spatially aligning the source images with the reference image. The essential problem in image alignment is to determine the appropriate mathematical model relating pixel coordinates in the source images to pixel coordinates in the reference image, and estimate the correct alignment among images with overlapping area, then panorama is created after choosing a final compositing surface for warping the aligned images.[10] Image alignment algorithm should not only ensure the accuracy of alignment, but also make the calculation not too large, which is a key step in the current image stitching technology.

2.4. *Image blending*

Once we have aligned all of the input images with respect to each other, we need to decide how to produce the final stitched mosaic image. Due to differences such as illumination, viewpoint, environment, image source, etc. in image acquisition, simply taking an average value at each pixel in the overlapping area of images will lead to visible seams, blurring and ghosting. In order to effectively minimize visible seam, image blending techniques have come to being.

Image blending is a kind of data fusion, it refers to the processing of images information that is analyzed automatically and combines optimally acquired image information under different conditions according to certain criteria by using computer image processing techniques, for the completion of the required decision and task assessment. Image blending algorithm should satisfy some characteristics. First, the method should be gradual, in order to ensure an invisible transition of the overlapping area; secondly, image blending process should be effective only on the overlapping area of the stitching, the information of the nonoverlapping area of the original images cannot be lost; finally, the complexity of the algorithm cannot be too high and affects the speed of the image stitching. Image blending techniques can be divided into three levels: pixel-level, feature-level, decision-level blending.

3. Algorithm Description

This paper presents a preprocessing method — histogram matching first; then SIFT algorithm is performed for image alignment; the weighted average blending method is used to remove the splicing gap and ultimately achieve seamless mosaic image.

3.1. *Image preprocessing*

In image preprocessing, the paper uses the histogram matching for the adjustment of images, so that the two images are at the same level of illumination.[11]

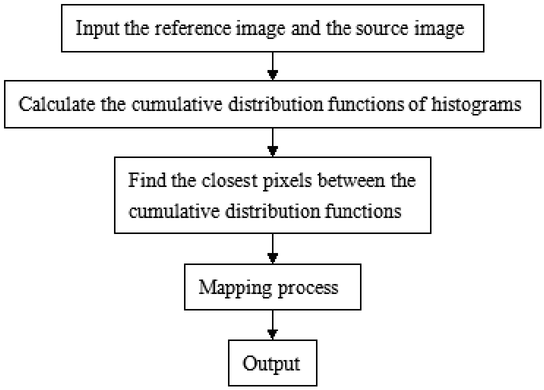


Fig. 1. The process of histogram matching.

Histogram matching is the method used to generate an image that has a specified histogram.[3] The process of histogram matching takes in an input image and produces an output image that is based upon a specified histogram. The required parameters for this algorithm are the input image and the specified image from which the specified histogram can be obtained. The steps of histogram matching are shown as Fig. 1.[11]

3.2. *Feature extraction based on SIFT*

The essence of the SIFT algorithm is to find the key points in different scale-space and calculate the orientation of the key points. The SIFT feature descriptor is invariant to image scaling, translation, rotation, and partially invariant to illumination changes and affine or 3D projection. Experimentally, the SIFT algorithm has been proven to be very useful in practice for image matching under real-world conditions. The major stages of the method include scale-space extrema detection, key point localization, orientation assignment, key point descriptor.[7]

3.2.1. *Scale-space extrema detection*

Scale-space theory is to convolve the original image with variable-scale Gaussian function for the construction of multiple levels of Gaussian pyramid, and scale-space extrema are detected in the constructed Gaussian pyramid. Gaussian convolution kernel is the only possible scale-space kernel for scale-change.

The scale-space of a two-dimensional image is produced by itself convolved with a Gaussian function $G(x, y, \sigma)$:

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y), \tag{1}$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}, \tag{2}$$

where $L(x, y, \sigma)$ is the output by Gaussian transformation, σ is the scale parameter, $I(x, y)$ is an input image, $*$ is the convolution operation in x and y.[2] When σ becomes larger, the overall profile of the image becomes clearer; when σ becomes smaller, the details of the image becomes better.

In order to efficiently detect the scale-space extrema, Lowe proposed convolving the difference-of-Gaussian (DOG) function with the image to get $D(x, y, \sigma)$[7]:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma). \qquad (3)$$

Once DOG images have been obtained, the key points are identified by comparing each pixel in the DOG images to its 26 neighbors in 33 regions — eight neighbors at the same scale and nine corresponding neighboring pixels in the adjacent scale above and below. It is selected as a key point candidate if it is the maximum or minimum among all compared pixels.

The key point candidates detected through the above method contain many low contrast points and some points which are poorly localized along an edge. A detailed fit is needed to perform for the location of the exact feature points and improve the anti-noise ability.

3.2.2. *Orientation assignment*

The key point descriptor must be invariance to image rotation, so we need to assign a consistent orientation to each detected key point. Considering distribution characteristics of the gradient orientations of the points within a region around the key point, the gradient magnitude and direction of the key point can be calculated as:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}, \qquad (4)$$

$$\theta(x,y) = \tan^{-1}\left[\frac{(L(x,y+1) - L(x,y-1))}{(L(x+1,y) - L(x-1,y))}\right], \qquad (5)$$

where L is the scale of the key point.[2] An orientation histogram is formed from the gradient orientations of sample points within a region around the key point. Peaks in the orientation histogram correspond to dominant directions of local gradients.[7]

Thus, by the method described above, each key point has three basic information: location, scale and orientation. The next step is to create a descriptor for the key point which is invariant to illumination changes or 3D viewpoint.

3.2.3. *Key point descriptor*

Before the computation of the key point descriptor, the coordinates of the descriptor and the gradient orientations should be rotated relative to the key point orientation first. Lowe proposed that take the feature points as the center, select 16×16 pixel neighborhoods and divide the region into eight 4×4 subregions. A $4 \times 4 \times 8 = 128$ element feature vector for each key point will be achieved by computing the orientation histogram with eight orientation bins in each subregion. Finally, the feature vector normalization is performed to reduce the effects of illumination change.[7]

3.3. *Rough matching process*

When the feature vectors have been generated, use Euclidean distance of the feature vectors to measure the similarity of the key point descriptors between two images; that is to say, select a key point in the reference image, calculate the Euclidean distance with its neighbor from the image to be stitched, and find the closest neighbor and second-closest neighbor by comparing the distance. Matches can be accepted for which the ratio between the distances of the closest and the second-closest is less than a certain threshold value.

Let feature descriptor be N-dimensional, the Euclidean distance of two feature descriptors is defined as:

$$d(i, j) = \sqrt{\sum_{m=1}^{N} (d_i(m) - d_j(m))^2}. \tag{6}$$

3.4. *Fine matching process*

The inaccurate matches can greatly affect the results of the model of the geometric transformation of the matching images. In order to improve the registration speed and accuracy, this paper uses random sample consensus (RANSAC) algorithm to discard the false matches. RANSAC is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers, and find the optimal fitting result. The algorithm was first published by Fischler and Bolles in 1981.[1]

The basic RANSAC algorithm in SIFT key point matching is summarized as follows: select RANSAC sample from sample set randomly, namely four matches; calculate the transformation matrix according to these four matches; test the current transformation matrix by asking how many of the remaining data matches are 'close' to the consensus set according to the sample set, the transformation matrix and the error metric function; determine whether the consensus set is optimal based on the current number of the consensus matches, optimum (maximum) consistent sets, if yes, update the current optimal consistent sets; update the current error probability, when it exceeds the minimum allowed error probability. Repeat Steps 1–4 until the current error probability is less than the minimum error probability.[4]

3.5. *Image blending*

Before we can register and align images, we need to establish the appropriate mathematical model relating pixel coordinates in one image to pixel coordinates in another. Transform the images according to the mathematical relationship, so that the registered images can be in the same coordinate. Finally, we need to develop algorithms to seamlessly blend the images with overlapping area. Therefore, finding the mathematical model to transform the registered images is a very important step.[6]

For two images to be stitched, the transformation between them includes translation, scaling, rotation, etc. The combination of these deformation forms several image transformation modes, and one of them which can satisfy all the image transformation models is called perspective transformation.

Suppose $I(x, y)$ and $I'(x, y)$ are the corresponding points on the reference image and the registered image, and their relationship can be described using the eight-parameter projection transformation (represented using homogeneous coordinates).

$$\begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \tag{7}$$

where $(x_i', y_i', 1)$ and $(x_i, y_i, 1)$ are the homogeneous coordinate representations of the ith point.[5] The transformation matrix H has eight degrees of freedom, in theory, only four pairs of the feature points can be used to estimate the transformation matrix H, by solving the linear equations, the parameters of the matrix H can be obtained, then normalize the pixels in the registered image and map them to the reference image coordinate system.

Once the images have been registered, the next step is to produce the final mosaic image. Because of the presence of illumination change, parallax, lens distortion, scene motion and exposure differences, stitching two images could create a visible seam. High-quality techniques for optimal seam finding and blending are an important component of image stitching systems.[8]

The selection of the blending strategy should meet two requirements: remove the splicing gap and achieve seamless photo-mosaics; try to ensure the original image information without loss. Many methods exist to perform image blending. Some well-known image blending methods are averaging, principal component analysis (PCA)-based image blending, wavelet transform image blending, pyramid and gradient domain blending, etc. In this paper, a simple weighted average blending method is used for gap closing.

The main idea of the algorithm is that the gray values of the image pixels in the overlapping area are obtained by weighted average gray values of the corresponding points in two images,[12]

$$I_{12}(x, y) = \begin{cases} I_1(x, y), & (x, y) \in S_1, \\ k(I_1(x, y)) + (1 - k)(I_2(x, y)), & (x, y) \in S_{12}, \\ I_2(x, y), & (x, y) \in S_2, \end{cases} \tag{8}$$

where I_1 represents the reference image and I_2 represents the registered image, I_{12} represents the blended image, S_1 represents nonoverlapping area in image 1, S_2 represents nonoverlapping area in image 2, S_{12} represents the overlapping area of two images. k is an adjustable factor and typically a number between 0 and 1, i.e. in the overlapped area, k is changed gradually from 1 to 0 along the direction from image 1 to image 2 so as to realize the seamless mosaic in overlapping area, as in Fig. 2.

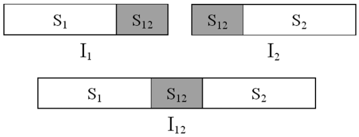


Fig. 2. Weighted average blending method.

Normally $k = d_1/(d_1 + d_2)$, d_1 and d_2 express, respectively, the distance from the point in the overlapping area to the left edge and the right edge of S_{12}.

4. Experimental Results and Analysis

To validate the image stitching algorithm proposed in this paper, a lot of experiments have been implemented on the computer of Intel (R) Core(TM) i5-4210U CPU @1.70 GHz, 2.3 9 GHz, 4.00 GB RAM, via MATLAB R2009a.

In this paper, three sets of images with overlapping area taken in the outdoor are as below for stitching. All images are normal exposure, except that one image of

(a) Image 1 (b) Image 2

Fig. 3. The first set of images (normal exposure).

(a) Image 3 (b) Image 2

Fig. 4. The second set of images (one image is overexposed).

the second set is overexposed and one image of the third set is underexposed (see Figs. 3–5).

Two different experiments were done in each set. In the first experiment, we use the SIFT algorithm to extract the key points of the two images and perform the rough matching process, and RANSAC algorithm is adopted to discard the false matches and improve the registration speed and accuracy. The mapping relationship between two images is obtained by using the fine matches. Then a simple weighted average method is used for image blending. The difference between the second experiment and the first is adding an additional step which is using the histogram

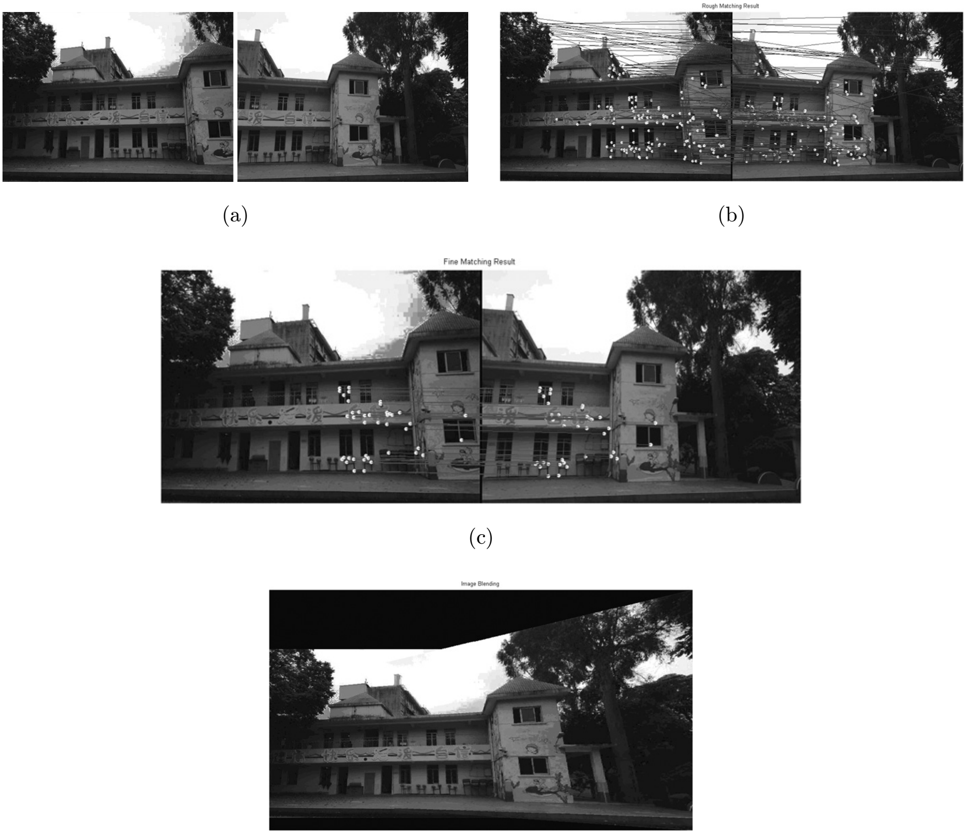
(a) Image 4 (b) Image 2

Fig. 5. The third set of images (one image is underexposed).

(a) (b)

(c) (d)

Fig. 6. First experiment: image stitching for the first set of images. (a) Images 1 and 2, (b) SIFT matches, (c) RANSAC inliers, (d) rendered with weighted average blending.



(a) (b)

(c)

(d)

Fig. 7. Second experiment: image stitching for the first set of images. (a) Histogram matching for images 1 and 2, (b) SIFT matches, (c) RANSAC inliers, (d) rendered with weighted average blending.

matching for the adjustment of images before feature extraction, so that the two images are at the same level of illumination.

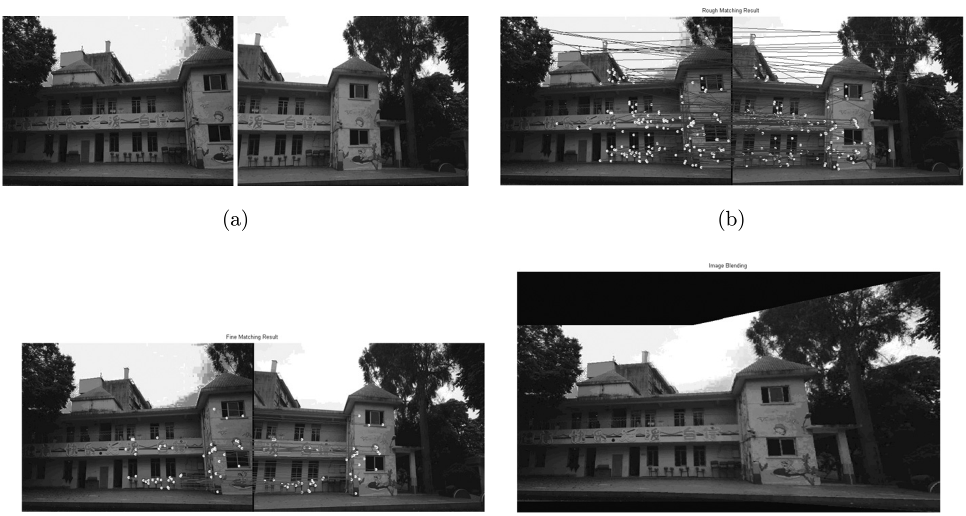
Figures 6–11 show the process and result of the two experiments in each set of images. From the quality of the final mosaic images, we can see that for the images

(a) (b)

Fig. 8. First experiment: image stitching for the second set of images. (a) Images 2 and 3, (b) SIFT matches, (c) RANSAC inliers, (d) rendered with weighted average blending.

(c) (d)

Fig. 8. (*Continued*)



(a) (b)

(c) (d)

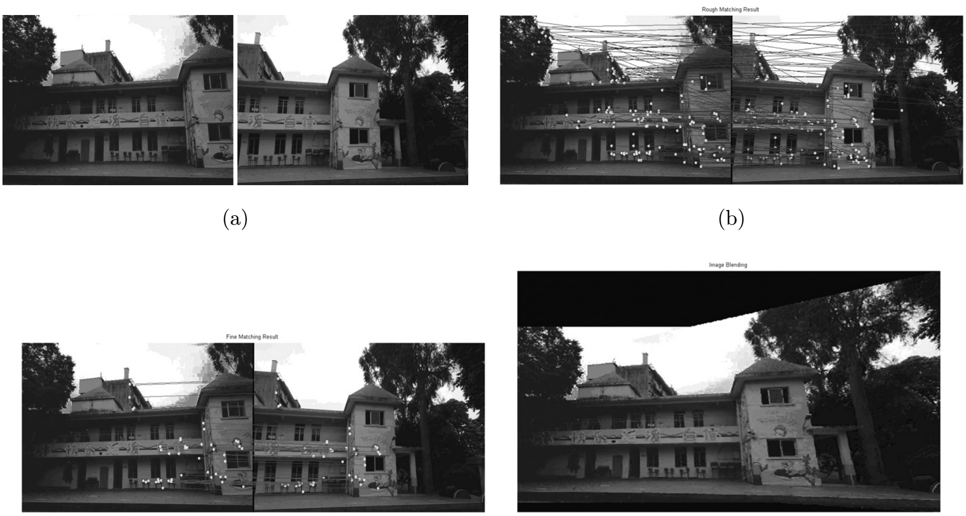
Fig. 9. Second experiment: image stitching for the second set of images. (a) Histogram matching for images 2 and 3, (b) SIFT matches, (c) RANSAC inliers, (d) rendered with weighted average blending.

(a) (b)

Fig. 10. First experiment: image stitching for the third set of images. (a) Images 2 and 4, (b) SIFT matches, (c) RANSAC inliers, (d) rendered with weighted average blending.

(c) (d)

Fig. 10. (*Continued*)



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

Fig. 11. Second experiment: image stitching for the third set of images. (a) Histogram matching for images 2 and 4, (b) SIFT matches, (c) RANSAC inliers, (d) rendered with weighted average blending.

with small difference in light intensity, the quality of the mosaic image is satisfactory if SIFT algorithm is performed directly (Figs. 6 and 7). Comparing Fig. 8 with Fig. 9, and Fig. 10 with Fig. 11, the mosaic images of Figs. 9 and 11 look more natural, so images which have significant illumination changes are preprocessed first by histogram matching that could improve effectively the accuracy of matches and optimize the mosaic image.

5. Conclusion

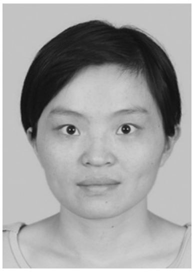
Image stitching algorithm has been the hot topic in the field of computer vision, it has a wide range of applications in 3D modeling, virtual reality, medical image processing, and other fields. This paper presents an image stitching algorithm based on histogram matching and SIFT algorithm. It not only retains the strong matching ability of classical SIFT algorithm, even though the presence of complex situations such as scaling, rotation, affine, the original characteristics of stable matching can be retained, but also preprocesses by histogram matching to reduce the matching errors and improve the quality of image stitching. Experimental results show that for the images that have significant illumination changes, the visual effect of using this image stitching method is good.

Acknowledgements

The authors would like to thank the Guangdong Provincial Science and Technology Project (2013B090600060) for financial support.

References

1. M. A. Fischler and R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* **24**(6) (1981) 381–395.
2. Q. Gao, H. Y. Yang and W. Yang, An image matching algorithm based on difference measure and improved SIFT algorithm, *J. Inf. Comput. Sci.* **11**(10) (2014) 3631–3642.
3. R. C. Gonzalez, R. E. Woods and S. L. Eddins, *Digital Image Processing Using MATLAB*, 2nd edn. (Publishing House of Electronics Industry, Beijing, 2005).
4. W. J. Huang and X. W. Han, An improved RANSAC algorithm of color image stitching, in *Proc. 2013 Chinese Intelligent Automation Conf.* (Springer-Verlag Berlin Heidelberg, 2013), pp. 21–28.
5. X. H. Jiang *et al.*, A steady image stitching algorithm based on SIFT and improved RANSAC, *Comput. Knowl. Technol.* **11** (2015) 127–132.
6. C. W.Kim, H. S. Sung and H. J. Yoon, Camera image stitching using feature points and geometrical features, *Int. J. Multimedia Ubiquitous Eng.* **10**(12) (2015) 289–296.
7. D. G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* **60**(2) (2004) 91–110.
8. Z.Qu, S. P. Lin, F. R. Ju and L. Liu, The improved algorithm of fast panorama stitching for image sequence and reducing the distortion errors, *Math. Probl. Eng.* **2015** (2015) 428076, 12 p.
9. X. B. Shi, X. S. Huang and D. Y. Zhang, Fast stitch algorithm on aerial images, in *Proc. 2013 Int. Conf. on Mechatronic Sciences, Electric Engineering and Computer* (IEEE, Shenyang, China, 2013), pp. 2213–2217.
10. R. Szeliski, *Computer Vision: Algorithms and Applications* (Springer, London, 2011).
11. J. Zhang and G. X. Chen, Research on the color correction algorithm of images based on histogram matching, *Appl. Mech. Mater.* **469** (2014) 256–259.
12. Y. Zuo, Y. Chen and H. You, A fast SIFT image mosaic algorithm based on wavelet transformation, *J. Chongqing Normal Univ., Nat. Sci.* **31**(3) (2014) 77–81.



Jing Zhang received her B.E. degree and M.E. degree in Printing Engineering from the School of Printing and Packaging, Wuhan University, China in 2001 and 2005. She is currently a PhD student at the School of Light Industry and Engineering, South China University of Technology, China. Her research interests include computer vision, computer graphics, computational geometry and image processing.

Guangxue Chen is a Professor and Doctoral supervisor at South China University of Technology, Guangzhou, China. He is also the team leader of digital printing and green packaging of State Key Laboratory of Pulp & Paper Engineering. His primary research interests include the printability of the paper, digital printing, color and media technology, new materials of printing and packaging. He undertook more than 20 research projects at national and ministry level and has published more than 200 papers and 60 of them were retrieved by SCI/EI.

Zhaoyang Jia received his B.E. degree in 2002 and his M.E. degree in Printing Engineering in 2008 from Tianjin University of Science & Technology, Tianjin. He is currently a PhD student at the School of Light Industry and Engineering of South China University of Technology, Guangzhou. He was a Lecturer in the College of Packaging and Printing Engineering of TUST, Tianjin. His research interests include color duplicate based on spectral image, print color quality control as well as computer graphics.