Journal of Korea Multimedia Society Vol. 14, No. 2, February 2011(pp. 201-209)

SURF와 멀티밴드 블렌딩에 기반한 파노라마 스티칭

라 연†, 신성식††, 박현주†††, 권오봉††††

요 약

이 논문은 이미지 매칭 알고리즘의 일종인 수정된 SURF(Speeded Up Robust Feature)와 이미지 블렌딩 알고리즘의 일종인 멀티밴드 블렌딩으로 구성된 파노라마 이미지 스티칭 시스템을 제안한다. 이 논문은 처음에 수정된 SURF를 기술하고 SIFT(Scale Invariant Feature Transform)와 비교하여 SURF를 이 시스템에서 채택한 이유에 대하여 논한다. 그리고 멀티밴드 블렌딩에 대하여 기술하고, 이어서 제안된 파노라마 이미지 스티칭 시스템의 구조에 대하여 설명하고 마지막으로 이미지 질과 처리시간에 대한 평가를 한다. 평가결과는 제안된 시스템이 개별 이미지들을 이음매 없이 연결하였으며, 많은 개개의 이미지 데이터에 대해서도 완전한 파노라마 이미지를 생성하였으며 처리 시간도 SIFT보다 빨랐다.

Stitcing for Panorama based on SURF and Multi-band Blending

Luo Juan†, Sungsik Shin††, Hyun Ju Park†††, Ou-Bong Gwun††††

ABSTRACT

This paper suggests a panorama image stitching system which consists of an image matching algorithm: modified SURF (Speeded Up Robust Feature) and an image blending algorithm: multi-band blending. In this paper, first, Modified SURF is described and SURF is compared with SIFT (Scale Invariant Feature Transform), which also gives the reason why modified SURF is chosen instead of SIFT. Then, multi-band blending is described. Lastly, the structure of a panorama image stitching system is suggested and evaluated by experiments, which includes stitching quality test and time cost experiment. According to the experiments, the proposed system can make the stitching seam invisible and get a perfect panorama for large image data. In addition, it is faster than the sift based stitching system.

Key words: Panorama(파노라마), SURF(서프), Stitching(스티칭), Multi-Band Blending(다중밴드블렌딩), LM(엘엠), Bundle Adjustment(번들조정)

1. INTRODUCTION

Stitching multiple images together can create beautiful high-resolution panoramas. It is a popular method of effectively increasing the field of view of a camera, by allowing several views of a scene to be combined into a single view, which is one of the most popular applications of image registration and blending. Image stitching algorithms have been used to produce digital maps and satellite photos, and it comes with today's digital cameras, there are more applications, such as video stitch-

※ 교신저자(Corresponding Author) : 권오봉, 주소 : 전북 덕진구 덕진동 1가 전북대학교 공대7호관 601호 (561-756), 전화 : 063)270-3402, FAX : 031)229-8281, E-mail : obgwun@jbnu.ac.kr
접수일 : 2010년 10월 11일, 수정일 : 2010년 11월 24일
완료일 : 2010년 12월 30일
† 학생회원, 전북대학교 컴퓨터공학
(E-mail : quihehappy@hotmail.com)
†† 학생회원, 전북대학교 컴퓨터공학
(E-mail : gshinnom@gmail.com)
††† 학생회원, 전북대학교 컴퓨터공학
(E-mail : loveofhj@gmail.com)
†††† 전북대학교 컴퓨터공학부 교수

ing, 3D image stitching, etc.

In general, stitching includes two main parts: image matching and image blending. Image matching is used to find the motion relationship between two images or several images, and it directly relates to the success rate and the speed of image stitching process. So if there is a high speed matching algorithm with high quality, it is better to choose the high speed matching algorithm. There are many algorithms for matching, and there are several kinds blending algorithms. For matching, there are two ways: direct method and feature detection method. Direct method is sometimes inconvenient because it always needs a high quality image. SURF and SIFT are two of the most famous feature detection methods because of their robustness. SIFT was presented by Lowe in 2004[1], which is invariant to scale changes and rotation. So it was widely used in object recognition, registration, stitching, etc. There are also some researchers present some new methods based on SIFT by giving some improvement. Ke and Sukthankar [2] used PCA(Principle Component Analysis) to normalize gradient patch instead of histograms. They showed that PCA-based local descriptors were also distinctive and robust to image deformations. In 2006, Bay and Tuytelaars [3] speeded up robust feature[SURF] detection algorithm and used integral images for image convolutions and Fast-Hessian detector, which created a new tide of robust feature detection method. References [4,5] are two applications of SIFT in Panorama. There are also some researches about comparing the famous feature detection methods [6-8].

The other part of stitching is blending. But before blending, what we need to do is to adjust the images into a same coordinate, which can be implemented by homography matrix that computed by matches that got previously. Bundle adjustment is a popular way to adjust images. In Brown and Lowe's panorama recognition [5], they used to unify the various illumination of the adjacent images and eliminate inconsecutive color which caused by the geometric correction or dynamic scene.

There are some blending algorithms: weight averaged linear blending, multi-band blending, gradient domain blending, etc. This paper chooses multi-band blending method because of its good performance, which can be easily got by experiment also [9], though it is a compromise between quality of result and time cost.

In short, we propose a panorama image stitching process which combines an image matching system(Modified SURF) and an image blending algorithm(multi-band blending). The process of our method is as follows. In the first stage, we acquire the SURF features from the image, then find the correct matches using KNN(K-nearest neighbor) and RANSAC(Random Sample Consensus), finally estimate the homography matrix according to LM (Levenberg-Marquardt) method[10]. In the second stage, we adjust the coordinate of images based on homography matrix. In the third stage, we blend the images by multi-blending to remove the stitch seam and illumination discrepancy [11].

This paper is structured in this way: Section 2 will represent the stitching process of this paper, and give description of modified SURF, bundle adjustment and multi-band blending method. Section 3 will give the experiment results of present stitching process. Section 4 will discuss some conclusion and future works. The images used in this paper are shown in Appendix.

2. STITCHING PROCESS

This section presents the process of image stitching of present system. There are two main parts. The first part is feature detection. The second part is image adjustment and image blending.

2.1 Robust Feature Detection

This part explores pre-step of our panorama image stitching: feature detection. First, this section gives a brief remark on SIFT and SURF, then shows that Modified SURF is more suitable than SIFT for our image stitching process based on an experiment.

2.1.1 Modified SURF

Comparing with SIFT, SURF algorithms employ slightly different ways of detecting features. SIFT builds an image pyramids, filtering each layer with Gaussians while increasing their sigma values and taking the difference. On the other hand, SURF creates a "stack" without 2:1 down sampling for higher levels in the pyramid, resulting in images of the same resolution[3]. Due to the use of integral images, SURF filters the stack using a box filter approximation of second-order Gaussian partial derivatives.

Integral images allow the computation of rectangular box filters in near constant time [3]. Present paper uses KNN to find the nearest neighbor with setting k to 2. In other papers, RANSAC [12] is used to estimate a model for finding the minimize error matches set, which can maintain the correct matches by comparing the distance of the closest neighbor to that of second-closest neighbor [1]. If it is less than the distance ratio, then it is maintained, else it is removed. Present paper decided to choose 0.5 as the distance ratio according to Lowe's experiment in SIFT. This makes the repeatability larger, which means improving of the matching performance. More details of this algorithm can be seen on [1,5].

2.1.2 Comparision of SIFT and Modified SURF

For present panorama stitching system, we evaluated SIFT and our Modified SURF with RANSAC [6]. They are evaluated in terms of their ability to reduce the number of false matches in given match sets, while preserving the good matches, which are used in the next step of feature detection. After extracting invariant scale features, we get potential feature matches by using k-nearest neighbor method, and then remove the mismatches with RANSAC algorithm. It enables to realize image match in precise. This part shows some comparison results of the paper for completeness. They are the comparisons of time cost and other performance (number of correspondence pairs, repeatability) as to illumination and scale change. Comparison results can be seen in more detail on [6]. The experiment was performed with object dataset in appendix A, whose sizes are all 300 x 240 pixels. Though SIFT get more matches, Modified SURF is much faster than SIFT [3,6,7].

In scale and illumination change, we got Fig.1, Fig. 2 and Table 1, Fig.1 and Fig. 2 give the match results of Modified SURF and SIFT in scale and illumination change respectively.

More details can be seen on Table 1, it shows that SURF is more robust than SIFT in illumination change. In Table 1, repeatability is computed as a ratio between the number of point-to-point correspondences that can be established for detected points and the mean number of points detected in two images using Eq. (1):

$$r_{1,2} = \frac{C(I_1, I_2)}{mean(m_1, m_2)} \tag{1}$$

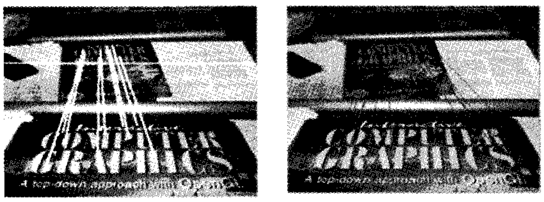


Fig. 1. Match results of modified SURF (left) and SIFT(right) in scale change.

Fig. 2. Match results of modified SURF (left) and SIFT (right) in illumination change.

Table 1. Repeatability in illumination Change

Object dataset image number	SIFT (repeatability)	SURF (repeatability)
11-12	43%	70%
11-13	32%	49%
11-14	18%	25%
11-15	8%	6%
11-16	2%	5%
average	21%	31%

, where C(I1,I2) is the number of the corresponding matching couples, and m1 and m2 are the numbers of the detected point respectively. Reference [3] showed that, though modified SURF was not better than SIFT in rotation, Modified SURF is as robust as SIFT in other performance. In panorama image stitching, there is not very large rapid rotation in general, so this paper chooses modified SURF to be the feature detection method weighted on its time cost and its good performance in scale and illumination.

2.2 Image Adjustment and Blending

This part explores post - steps: image adjustment and blending for panorama image stitching. It describes the selected algorithms of image adjustment and blending and the reason why the algorithms are selected.

2.2.1. Image Adjustment

Image adjustment is to transform the images into a same coordinate or computing surface. In present paper, we choose bundle adjustment[13] as image adjustment algorithm because it is popular in computer vision and shows a good performance in Demo Software: SIFT Key point Detection [14].

The process of bundle adjustment is as follows: first, choose one of the images to be reference surface, then, transform each of other images to the reference surface, at the end of which all images are on the same surface. However, for larger fields of view, we cannot maintain a flat representation without excessively stretching pixels near the border of the image. The usual choice for compositing larger panoramas is to use a cylindrical or spherical projection.

Transformation needs to compute the homography and optimize the parameters of the matrix in the adjustment. The process is as follows: first find out the best neighbor image for each image, then directly calculate the distance between the two neighbor images, lastly minimize the distance value to adjust the matrix between the neighbor images. LM, which is one of the minimization methods using nonlinear minimum square evaluation, was one of the most popular algorithms for optimizing in bundle adjustment. In one sentence, what needs to be done is to minimize the transfer error, which is calculated as Eq. (2):

$$D=\sum(d(X_i,H^{-1}X_i')^2+d(X_i',HX_i)^2) \tag{2}$$

, where d is the Euclidian distance, xi and xi' are the corresponding points, the estimated homography H will be the one for which Eq. (2) is minimized, H-1 is the inverse of H. LM has become a standard technique for nonlinear least-squares problems and can be thought of as a combination of steepest descent and the Gauss-Newton method [10]. There is a free LM implementation that can be found at [15].

2.2.2 Image Blending

Mathematically speaking, image blending is nothing more than a type of interpolation. The goal of the blending is to produce a resulting image where no transition can be seen between the original source images. Average weighted blending is a common linear method, which employs simple average of images in the overlapping regions. This results in ghosting artifacts, blurring, and visible seams that degrade the mosaic, so as bilinear blending. But Linear blending method is fast and can be a good compromise between quality and speed if you are not too demanding on quality. The work of Burt and Adelson[11] on multi-band blending or called by pyramid blending has proved particularly effective for image stitching without blurring and ghosting artifacts. It will produce much better results than the "Linear" mode. The transitional areas between images will hardly visible. The comparison results can be seen directly in [9,14]. Multi-band blending scheme ensures smooth transitions between images despite illumination differences, while preserving high frequency details. So we choose multi-band blending method (2-band) in this paper.

The idea of multi-band blending is to blend low frequencies over a large spatial range and high frequencies over a short range. Image is decomposed into a collection of N band pass images using the Laplacian pyramid. The Laplacian pyramid of the final image is formed as Eq. (3):

$$Y_k(i,j) = X_{1,k}(i,j)M_k(i,j) + X_{2,k}(i,j)(1 - M_k(i,j)) \quad (3)$$

where X1,k and X2,k are the kth level of Laplacian pyramid decomposition for the two images after coordinate adjustment, Yk is the kth level of Laplacian pyramid decomposition for the final combination result. And Mk is the kth level of Gaussian pyramid decomposition of the image mask. Pyramid blending gradually blends the lower frequencies of the images while maintaining a sharper transition for the higher frequencies, which makes image clearer than other methods. The results of experiment will show in the follow section.

3. EXPERIMENTAL RESULTS

The experiment has been done on visual studio 2008 C++ environment with OpenCV library. Before giving the experimental results, we show the pseudo code of our stitching system in Fig. 3.

3.1 Process of present system

There are two main parts; matching and blending. The connection of the two parts is the correspondence pairs. In the following, there will be more details of the stitching algorithm.

In matching part,

- First, Modified SURF is used to detect feature points. Modified SURF feature descriptors are 64 dimensional vectors in this experiment.
- Secondly, among this descriptors, square Euclidean distance ratio between neighbors is calculated, which is computed by the closest neighbor to the second closest neighbor. If it exists and its value is lower than a threshold value (setting as 0.5), It is maintained as neighbors. All of which can be described by K-NN(2).
- Lastly, RANSAC is used to estimate a model of consensus set that minimizes matching error. This is an iterative process, the goal of which is to find the largest feature points good to transformation. The matching pairs fitted to the minimizing error model are known as the correct matches or correspondences.

In blending part,

- Homography matrix H can be estimated (Eq.(2)) based on correspondence pairs which got from previous step, which can be computed by LM [12, 15].
- Then adjustment between images is done, in other word, the images has been transformed into the corresponding image in a same coordinate system by the H matrix. Now the panorama has already got, but with some stitching error of color and illumination.
- Lastly, Multi-band blending is used to make the transformation error or the seam invisible by resetting the value of overlapped region according to Eq.(3) [16].

According to the previous explanation, the Pseudo code of this stitching algorithm is shown in the Fig. 3.

3.2 Experiments

Experiments consist of two parts: panorama

```
Input image I1....In
for   i→1 to n   do
   Detect SURF features;
for   i→1 to n
   for   j→1 to n   do
      if(i!=j)
         Check all the SURF features by KNN;
            if [dis(closet)/dis(secdloset)] less than ratio
               the two features is a match pair;
Output are match pairs;
for all the match pairs   do
   Compute Homography matrix H;
Transform all Images by H to a same surface;
for   previous result   do
   2-band blending;
Output image is panorama image.
```

Fig. 3. The pseudo code of present stitching system.

quality (stitching) test and time cost experiment. A good stitching program should make panorama seamless and clear and be fast for using in various application such as real time image processing. In this experiment, we use datasets as shown in appendix A, B [17].

3.2.1 Stitching Test

Fig. 4 and Fig. 5 show the results of stitching before multi-band blending and after multi-bland blending. For those image, we use 4 images; img11, img12, img13, img14 from Road dataset as shown in Appendix B. It is easy to find out the three seams on the Fig. 4, which caused by the illumination and color. Fig. 5 looks very clear and perfect.

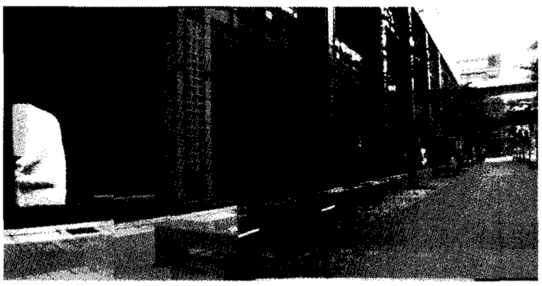


Fig. 4. Panorama with obvious seams(No blending).

Fig. 5. Panorama with seamless(with blending on Fig. 4).

Next, we will do an experiment with large data set. In this experiment, we use 16 images of Road dataset. Fig. 7 is the result of present program and Fig. 6 is the result of SIFT system [14].

The present stitching system can also show its good performance on large image dataset. The present stitching system shows as good performance as SIFT in Fig. 6. Fig. 8. shows another stitching results of present system by Grass data set consisting of 17 images

3.2.2 Time Cost Experiment

Due to using the fast matching method: modified SURF, present system is much faster than SIFT

Fig. 6. Panorama stitched 16 images based on SIFT.

Fig. 7. Panorama stitched 16 images based on modified SURF.

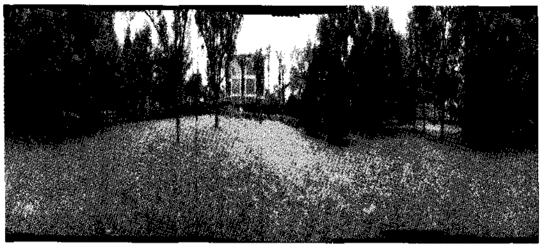


Fig. 8. Panorama stitched 17 images using grass data set.

Table 2. Time cost between SIFT and present system

Image data set	Road	Bridge	Camp
SIFT demo(ms)	24949.51	19478.04	22775.55
Present system(ms)	8271.54	8517.45	19872.29

demo as shown in Table 2. The processing time of SIFT demo is got by the performance tool of Visual Studio 2008. The data set that used in the experiments, all include 16 images. Camp data and Bridge data can be downloaded from [17].

4. CONCLUSION AND FUTURE WORKS

According to the Experiments, present system shows its good performance with seamless panorama and fast processing time even for a large image dataset. We can reason out the answer as follows. Correct matches can be found in case there are some changes of illumination, color, blur, etc using Modified SURF. Also its high speed contributes to shorting the processing time. And bundle adjustment and multi-band blending make the panorama seamless. In addition, LM is used to estimate the homography, which makes the transformation more accurate.

But the present system shows its defects when there are some noise images that are not neighbored. So as a future work, we plan to do more research about removing the noise before stitching processing. Also we consider panorama based virtual gallery[18] as a next research topic.

APPENDIX

The test images of present paper are shown in this appendix, which include two image dataset. They are Object image dataset and Road image dataset.

A. Object image dataset

The following images are Object image dataset. Images marked from left to right, from up to down is: 1, 2, 3, …, 16. All the images are in the same size with 574x 768.

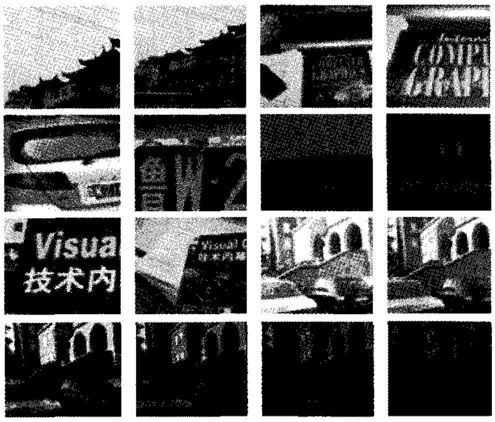


Fig. 9. Object image dataset.

B. Road image dataset

The following images are Road image dataset.

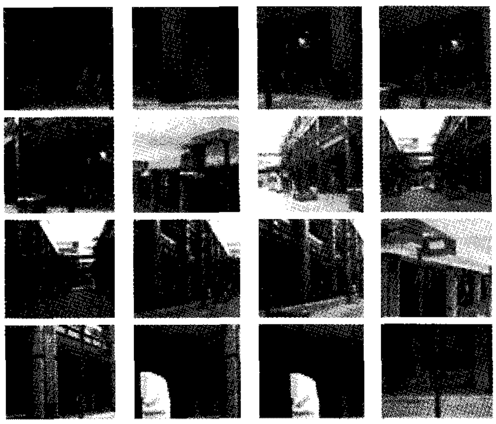


Fig.10. Road image dataset.

Images are marked from left to right, from up to down as 0, 1, 2.. 15. All the images are in the same size with 574x 768 resolutions.

REFERENCES

[1] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," IJCV, Vol.60, No. 2, pp. 91-110, 2004.

[2] Y. Ke and R. Sukthankar. "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors," Proc. Conf. Computer Vision and Pattern Recognition, pp. 511-517, 2004.

[3] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool, "SURF: Speeded Up Robust Features," Computer Vision and Image Understanding (CVIU), Vol.110, No.3, pp. 346-359, 2008.

[4] Zhan-long Yang and bao-long Guo, "Image Mosaic Based On SIFT," International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 1422-1425, 2008.

[5] M. Brown and D. Lowe, "Recognizing Panoramas," Proc. Ninth Int'l Conf. Computer Vision, pp. 1218-1227, 2003.

[6] Luo Juan and Oubong Gwun, "A Comparison of SIFT, PCA-SIFT and SURF," *International Journal of Image Processing* (IJIP), Vol.3, No.4, 2009.

[7] K. Mikolajzyk and C. Schmid, "A Perforance Evaluation of Local Descriptors," *IEEE, Trans. Pattern Analysis and Machine Intelligence*, Vol.27, No.10, pp. 1615-1630, 2005.

[8] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L.V. Gool, "A Comparison of Affine Region Detectors," IJCV, 65(1/2), pp. 43-72, 2005.

[9] "Interpolation and blenders," http://www.autopano.net/wiki-en/action/view/Interpolation_and_blenders.

[10] Manolis I. A. Lourakis, "A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar," Vol.11, Feb., 2005.

[11] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics.," *ACM Trans. On Graphics*, Vol.2, No.4, 1983.

[12] W. Zhang and J. Kosecka. "Generalized Ransac Framework for Relaxed Correspondence Prob lems," In International Symposium on 3D Data Processing, Visualization and Transmission, 2006.

[13] M. Brown and D. Lowe, "Automatic Panoramic Image Stitching using Invariant Features," International Journal of Computer Vision, Vol.74, No.1, pp. 59-73, 2007.

[14] "Demo Software," SIFT Keypoint Detection, http://people.cs.ubc.ca/~lowe/keypoints

[15] "A c/c++ library for robust, non linear homog raphy estimation," http://www.ics.forth.gr/~lourakis/levmar/.

[16] N. Mueller and T. Q., "NguyenImage. Inter Polation Using Classification and Stitching," ICIP08, pp. 901-904.

[17] "Image dataset," ftp://210.117.178.61/stitching image test dataset .

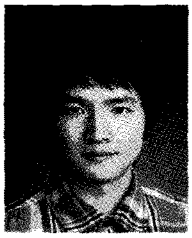
[18] Eung-kon Kim and Gyoung-nam Park, "Design and Implementation of a Web-based 3D Virtual Gallery using Panorama Images," 멀티미디어학회 논문지 제5권 제6호, pp. 655-664 2002.

라 연

2010년 8월 전북대학교 컴퓨터 공학과 석사 졸업
관심분야 : 컴퓨터그래픽스.

신 성 식

2006년 2월 전주대학교 컴퓨터 공학과 학사 졸업
2008년 2월 전북대학교 컴퓨터 공학과 석사 졸업
2008년 전북대학교 컴퓨터공학과 박사 재학
관심분야 : 3D컴퓨터그래픽스, 패턴인식, 소프트웨어 공학



박 현 주

2009년 2월 전북대학교 컴퓨터 공학과 학사 졸업
2011년 2월 전북대학교 컴퓨터 공학과 석사 졸업예정
관심분야 : 3D컴퓨터그래픽스, 소프트웨어 공학

권 오 봉

1980년 고려대학교 전기공학과 학사 졸업
1983년 고려대학교 전기공학과 석사 졸업
1993년 일본구주대학교 총합이 공학연구과 박사 졸업
1992년～1993년 일본구주대학 교 정보공학과 조수
1994년～현재 전북대학교 전자정보공학부 교수
관심분야 : 컴퓨터그래픽스, 비주얼라이제이션