MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11 - 13, 2002, Nara - ken New Public Hall, Nara, Japan

10—4

Super-Resolved Image Synthesis from Uncalibrated Camera with Unknown Motion

Shinichiro Sei*, Hideo Saito * †
Department of Information and Computer Science, Keio University

Abstract

In this paper, we propose a novel method for synthesizing super-resolved images from image sequence taken with uncalibrated moving camera by integrating all frames in the sequence based on projective geometry between the frames. In this method, assuming that the object scene can be approximated with multiple plane patches, all frames in the input image sequence are registered with the fundamental matrices between the frames, and then integrated by transforming with homographic matrices. For reducing the registration error, the blur mount in the integrated image is estimated so that the sub-pixel order registration between the frames can be achieved. For demonstrating the effectiveness of the proposed method, we confirm that the proposed method improves the resolution of the synthesized image comparing with the input image sequence.

1 Introduction

In this paper, we propose a new method for synthesizing super-resolved image from sequential images taken with uncalibrated camera with unknown motion. In this method, we assume that the objective scene can be represented by a number of planar patches. Then, each patches captured in all the input images are automatically registered and blended for synthesizing super-resolved image. In the proposed method, any camera parameter is not required in advance, but only fundamental matrices between the input images are required, which can be easily obtained by general feature-point tracking in image sequence [4].

As related works, Cheeseman et al.[1] proposed a method for super-resolution by blending sequential images with small displacement. They proposed a effective method for blending images based on Bayes' presumption, but they only assume that the object scene is a plane that is vertical to the camera direction. Irani et.al.[2] proposed a super-resolution based on affine camera model, but affine camera model requires limitations on the object scene. Mann et.al.[3] extended this method to perspective camera model, but the object scene must be a plane in this method. Our method can be regarded as an extended method of those studies, but the original contribution is proposing an efficient way to adapt a framework of the projective geometry to such super-resolution technique based on blending sequential images.

2 Proposed Method

The outline of this method is shown in Fig1. The input is a set of image sequence that is taken with uncalibrated motion camera. A super-resolved image is synthesized according to the following steps.

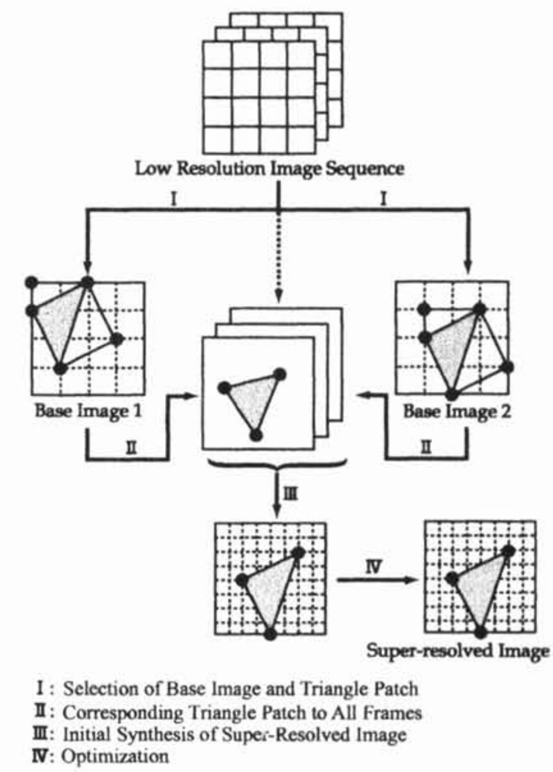


Figure 1: Outline of our method.

Selection of base images and triangle patch Two images are arbitrary chosen from the set of the images for defining projective 3D coordinate of the object space. The selected two images are called *base images*. In those base images, the object area is divided into some triangular patches, and then correspondence of the vertices between the base images are manually selected. Subsequent process is performed for every triangular patch defined here.

Corresponding the triangular patch to all frames Position of the vertices of the triangle patch in the base images are transferred into other images using fundamental matrices. In the proposed method, we assume that fundamental matrices between two base images and other images are previously estimated by feature point tracking of the input image sequence. According to the point to epipolar line correspondence, the vertex point on the two base images can be transferred into other images as two epipolar lines as shown in Fig.2. The intersection of the lines is corresponding point

*Address: 3-14-1 Hiyoshi Kohoku-ku, Yokohama 223-8522, Japan. E-mail: saito@ozawa.ics.keio.ac.jp
†"Information and Human Activity," PREST, JST, Japan

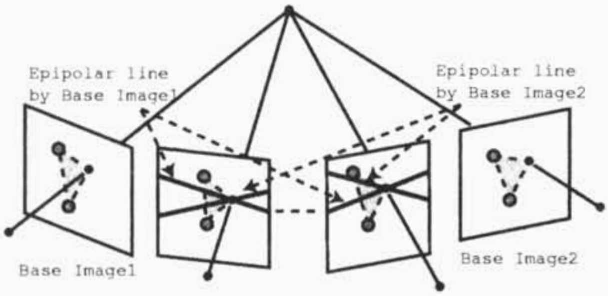


Figure 2: Calculating corresponding point

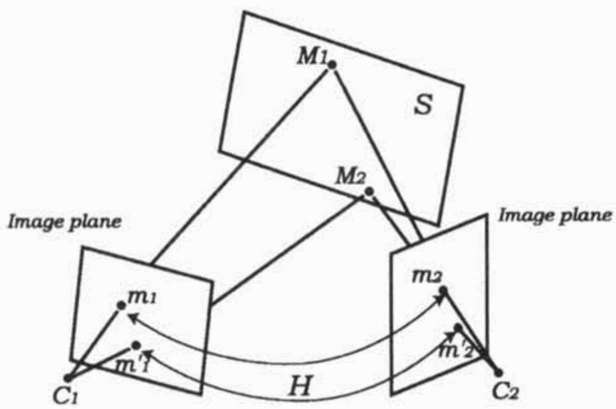


Figure 3: Homography

of the vertex point. In such a way, every vertex on every triangular patch on the base images can also transferred into all other images.

Then the correspondences of three vertices of a patch between images also provides homography H, which transfers all the pixels inside the patch between the images as shown in Fig.3. The homography H can be estimated from the three correspondences and the fundamental matrix F between the images.

Initial synthesis of super-resolved image Blending images on all triangular patches, higher resolution than the input images can be obtained, because the sampling position of pixel in each triangular patch is shifted in sub-pixel displacement. Such displacement enables to synthesize super-resolved image. We assume that each triangular patch can be approximated to be planer surface. According to this assumption, every triangle patch image can be transferred onto the same image plane by using the homography between the different view images. Such blending of all triangular patch images onto a same image plane, initial estimation of super-resolved image can be synthesized.

Optimization of the registration Even though the correspondence of the vertices between base images are detected manually, positioning error up to one pixel unit can not be avoided. Accordingly, the corresponding points to other images are not accurate. Such error in the correspondence makes the super-resolved image of initial synthesis degraded. For obtaining corresponding accuracy in sub-pixel order, we employ iterative optimization of the correspondence by evaluating quality of synthesized super-resolved image. In this optimization, the positions of the

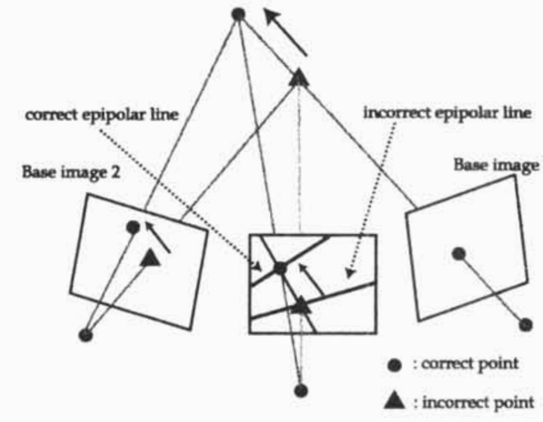


Figure 4: Optimization

corresponding vertices between two base images are regarded as objective variables. Then, the number of variables is 12, that is 3 vertices $\times$ 2 base images $\times$ (x, y) 2 variables. The 12-dimension vector that gives maximum quality is found by employing simplex method as optimizing algorithm.

Since the correspondence of the vertices points between the base images are transferred into other images by using epipolar geometry, the number of registered parameters is always 12 for each triangle patch, even if the number of sequential images is increased.

For the optimization, we need to evaluate the quality of the super-resolved image. In this paper, we use high-frequency energy of the super-resolved image. If the registration is not perfect, even same features in different images are blended at different position in the super-resolved image. Such blending of the features with a lag affects as blur in the super-resolved image. We assume that the blur can be measured by the amount of the high-frequency energy in the super-resolved image. For measuring the high frequency, we define the high-pass filter expressed as the following equation.

$$A(u,v) = 1 - G(u,v) = 1 - e^{\frac{u^2+v^2}{2}\sigma^2} \qquad (1)$$

This filter is applied to the super-resolved image $I_d(x,y)$, and then the filtered image $I_a(x,y)$ provides the evaluation value $S(p)$ for a triangular patch p.

$$\begin{aligned} I_a(x,y) &= a(x,y) * I_d(x,y) \\ &= I_d(x,y) - g(x,y) * I_d(x,y) \qquad (2) \\ S(p) &= \sum_{(x,y)\in p} I_a(x,y) \qquad (3) \end{aligned}$$

where $a(x,y)$ is a spatial representation of $A(u,v)$.

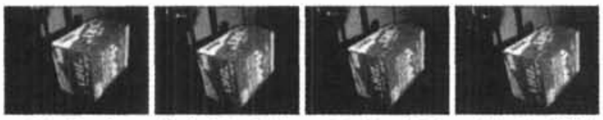
3 Experiments Results

We took image sequences by moving a camera by a hand. The captured sequential images are 160×120 pixels of color 24bit intensity resolution. The number of images is 200. By blending all sequential images, we extend the resolution into 640×480 pixels. Examples of the input images are shown in Fig.5 and Fig.6.

Fig.7 and Fig.9 show the comparison of the synthesized super-resolved image with an image in the input sequence. Fig.8 and Fig. 10 show zooming up of the images, where

(a) No.1 (b) No.100 (c) No.150 (d) No.200

Figure 5: Original image sequence (1)



(a) No.1 (b) No.100 (c) No.150 (d) No.200

Figure 6: Original image sequence (2)

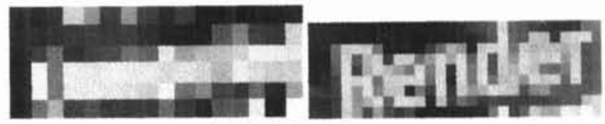
(a) shows an input image, (b) shows the bi-linear interpolated image of the input image, and (c) shows the super-resolved image synthesized by the proposed method. The letters, which are not readable in the input images (a) and (b), become readable in the super-resolved images (c). The super-resolution can be effectively achieved even for the scene with large depth variance as Fig.9, since the proposed method does not assume the orthographic or affine camera model, but performs super-resolution under the perspective camera model via epipolar geometry among the images.

In addition to the experiment with the actual camera demonstrated above, we perform an super-resolution experiment from input images that are virtually synthesized in the computer for evaluating the performance of the proposed method. First, we prepare a digital image with several times higher resolution than the input camera images by using an image scanner. Then, we virtually synthesized 200 input images with 160×120 pixels under various camera parameters as shown in Fig.11. Fig.12 shows example images of the synthesized image sequence. From the input images, we generate a super-resolved image with three times resolution of 480×360 pixels, and then evaluate the performance of the proposed method as follows.

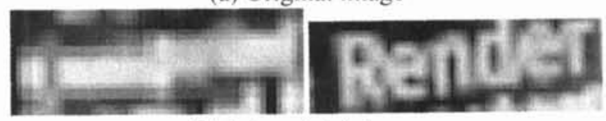
Performance of optimization Fig.13 and Fig.14 show the estimated vertices position of a triangle patch. In Fig.13, which is before the optimization, the vertices positions in (c) and (d) are different from the base images (a) and (b), because the initial estimate of the transferred triangle is affected by the error of the fundamental matrices. After the optimization, which is shown in Fig.14, the vertices positions in all images are almost same because of the effect to optimization for the registration.

(a) Original image (b) Super-resolved image

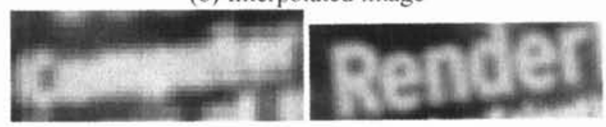
Figure 7: Super resolved images synthesized from real image sequence (1)



(a) Original image

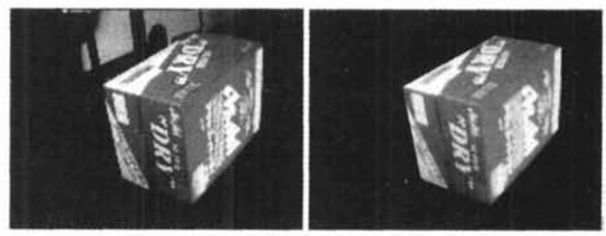


(b) Interpolated image



(c) Super-resolved image

Figure 8: Super resolved images synthesized from real image sequence (1) (zoom)



(a) Original image (b) Super-resolved image

Figure 9: Super resolved images synthesized from real image sequence (2)

Fig. 15 shows the comparison of the super-resolved images before optimization and after optimization. The blur caused by the registration error can be reduced by the optimization.

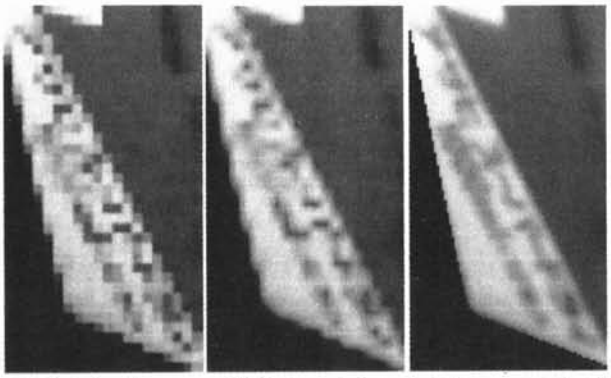
Quality of the super-resolved image We evaluate the quality of the super-resolved image by calculating SNR with virtually synthesized image with the same resolution of the super-resolved image under the scheme shown in Fig.11. SNR of the super-resolved image shown in Fig. 15 (b) is 21.6dB, while SNR of the interpolated image with bi-linear method is 20.1dB. Such improvement in SNR depends on the number of input images to generate the super-resolve image. Table 1 shows SNR of the super-resolved images generated from various number of input images. Those results demonstrate that the proposed method is valid to improve the quality of the input image.

4 Conclusion

We proposed a method of synthesizing super-resolved image from sequential images that are captured by moving

Table 1: SNR(dB) of super-resolved image and the number of input frames.

number of images	10	25	100	200
SNR (dB)	20.7	21.3	21.5	21.5



(a) Original image (b) Interpolated image (c) Super-resolved image

Figure 10: Super resolved images synthesized from real image sequence (2) (zooms)

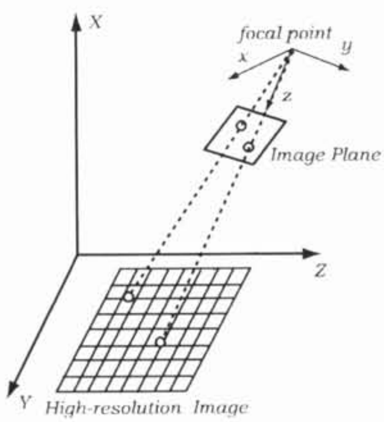


Figure 11: Synthesis of virtual motion camera image

camera with handy motion. The experiment results demonstrate that the proposed method is effective for obtaining higher resolution images than input images.

References

[1] P. Cheeseman, B. Kanefsky, R. Hanson, and J. Stutz. Super-resolved surface reconstruction form multiple images. *Technical Report FA-94-12 , NASA Ames Research Center , Artificial Intelligence Branch*, 1994.

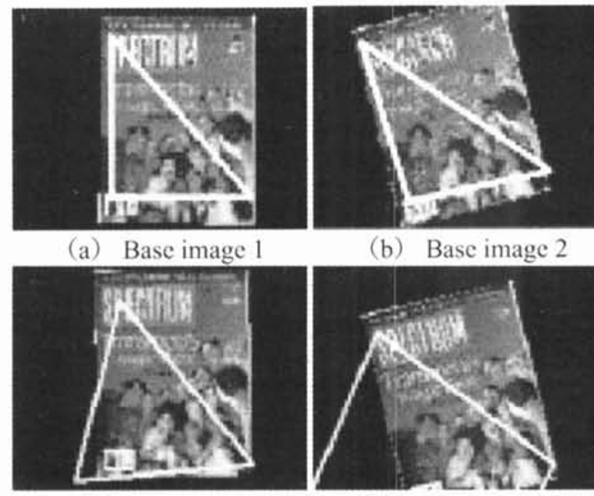
[2] M. Irani, and S. Peleg, "Improving reconstruction by image registration," CVGIP, vol.53, No.3, pp.231-239, 1991.

[3] S. Mann, and R. W. Picard, "Constructing high quality stills from video," Proc. ICIP'94, pp.363-367, 1994.

[4] Z. Zhang. Determining the epipolar geometry and its uncertainty : a review. *Research Report 2927 , INRIA Sophia-Antipolis , France*, July 1996.

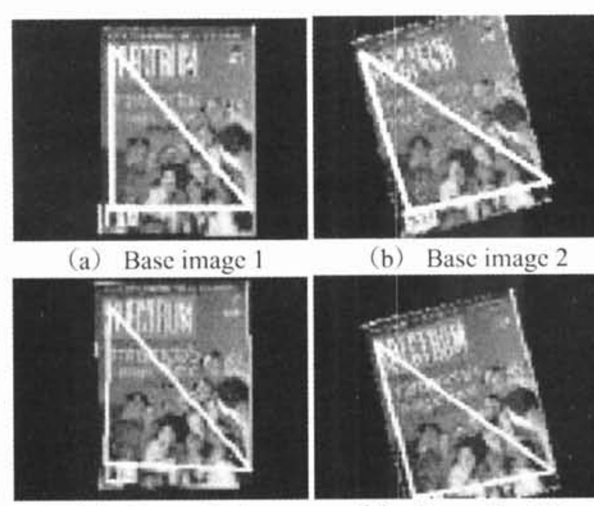
(a) No.1 (b) No.100 (c) No.150 (d) No.200

Figure 12: Synthesized Image Sequence



(a) Base image 1 (b) Base image 2
(c) Frame No.2 (d) Frame No.199

Figure 13: Result of calculating corresponding point (before optimization)



(a) Base image 1 (b) Base image 2
(c) Frame No.2 (d) Frame No.199

Figure 14: Result of calculating corresponding point (after optimization)

(a) Before optimization (a) After optimization

Figure 15: Comparison of super-resolved images before optimization and after optimization