Prototype of Super-Resolution Camera Array System

Daiki Hirao and Hitoshi Iyatomi(✉)

Applied Informatics, Graduate School of Science and Engineering,
Hosei University, Tokyo, Japan
iyatomi@hosei.ac.jp

Abstract. We present a prototype of a super-resolution camera array system. Since the proposed system consists of a number of low-cost camera devices, all of which operate synchronously, it is a low-cost, high quality imaging system, and capable of handling moving targets. However, when the targets are located near the system, parallax and differences in photographic conditions among the cameras become pronounced. In addition, conventional super-resolution techniques frequently emphasize noise, as well as edges, contours, and so on, when the number of the observed (i.e., low resolution) images is limited. Therefore, we also propose the following procedures for our camera-array system: (1) color calibration among cameras, (2) automated region of the interest (ROI) detection under large parallax, and (3) effective noise reduction with effective edge preservation. We developed a camera array system comprising 12 low-cost Web camera devices. We confirm that the proposed system in general reduces the drawbacks of the array system and achieves approximately a 2 dB higher S/N ratio, i.e., equivalent to the effect of two additional images.

1 Introduction

Super-resolution is a technique for generating a high-resolution (HR) image using observed low-resolution (LR) image (s), and many studies on these techniques have been reported [1]. The super-resolution techniques addressed in these research studies researches can be divided into two broad types. The first is the so-called "learning type" super-resolution technique. This type usually utilizes only one base image from the observed images and predicts unknown details using a pre-trained database and/or estimators [2] or an interpolation approach based on signal processing techniques [3]. These methods generate better quality images by estimating unknown high-frequency components in many cases; however, there is no guarantee that the estimations of the components are true.

The second type is the so-called "registration type" super-resolution technique. This registration-based super-resolution technique utilizes a large number of images in order to increase the pixel density of the image. Accordingly, the true high-frequency components of the image can be estimated by using an appropriate reconstruction algorithm. Typical registration type super-resolution techniques utilize a maximum likelihood (ML) method [4], maximum a

G. Bebis et al. (Eds.): ISVC 2015, Part I, LNCS 9474, pp. 911–920, 2015.
DOI: 10.1007/978-3-319-27857-5_81

posteriori (MAP) method [5,6], iterative backward projection (IBP) method [7], frequency domain approach [8], or projection onto convex sets (POCS) method [9,10] as a reconstruction algorithm. However, because these methods require not a few observed images, precise correction of the positional deviations among the images, namely sub-pixel image registration, is necessary for the successive reconstruction process. It should be noted that the positional deviation among the images is not only translation but also constitutes complex deformation, such as rotation and scaling.

Various image registration techniques, including region-based matching [11], local-feature-based matching based on robust image features such as SIFT [12] and SURF [13], and so on, have been widely used. In most cases, the super-resolution process is performed based on images taken by a single camera device and therefore when moving objects are present in the target image frame, sophisticated image processing techniques [14] are necessary.

In this paper, we propose a synchronously controlled super-resolution camera array system. This system is a high quality imaging system that can also handle moving objects and does not require expensive optical equipment or complex image processing. However, the camera array system may give rise to the following issues. (1) A mismatch of color and brightness in the images captured by different cameras, (2) parallax among cameras, and (3) limited image quality due to the restriction of the number of cameras. Thus, in this paper, we also propose three effective measures to address these issues.

2 Registration-Type Super-Resolution

Our super-resolution camera-array system is composed of a large number of low-cost cameras and a registration type super-resolution algorithm. The process of the super-resolution consists of two processes:

1. Image alignment process.
2. Reconstruction process.

The details are described in the following.

2.1 Image Alignment Process

The registration type super-resolution technique requires sub-pixel registration accuracy. A flow chart of the image alignment process is shown in Fig. 1. Now, let us consider the image alignment between two images, namely, the HR image estimated at the time of the imaging and the up-sampled observed LR image, for instance.

When an arbitrary point object is captured by the proposed camera array system, we assume it should be observed at a similar position in each image. Based on this assumption, we compared the coordinates of each matching point of the images and, if their coordinates were significantly different, we considered this matching point inappropriate, and therefore, eliminated it. According to the results of preliminary experiments, we eliminated 50 % of matching points. Then, we selected four

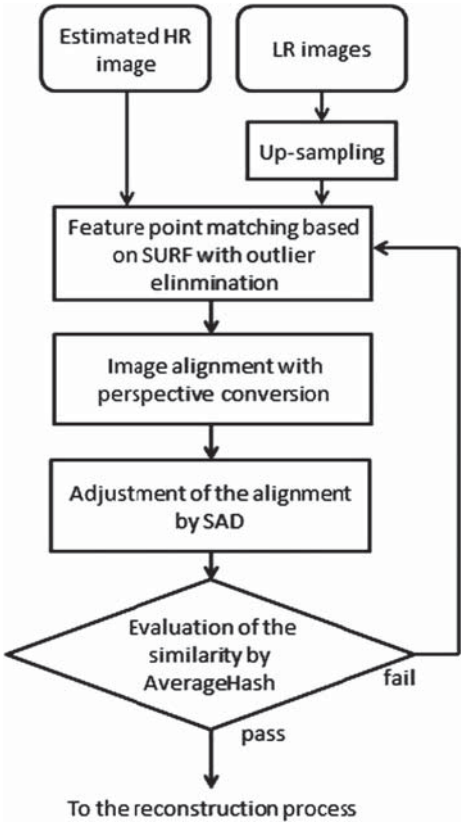

Fig. 1. Flow of the image alignment process.

matching points between the images for performing image alignment using perspective projection transformation. In order to achieve accurate image alignment, it is more effective to select four points having a bounded area that is as large as possible. Accordingly, we selected the matching point from each image quadrant that was the shortest distance from the corresponding corner. In order to achieve accurate and reliable image alignment, we conducted further image registration using the sum of absolute difference (SAD) measure. We generated several shifted images from one of the images and sought the location on the second image that had the lowest SAD. Thus, we realized a sub-pixel precise image alignment, i.e., alignment is performed in up-sampled images. Furthermore, in order to ensure accurate alignment, we evaluated the accuracy of the image alignment process. When we determined that the image alignment was inappropriate, we re-performed the alignment process using a different selection of SURF matching points. Here, we applied the AverageHash algorithm [15] for the evaluation of the alignment by calculating the similarity of the images. This algorithm first defines the region of interest (ROI) in the image and down-samples it to an 8×8 gray scale image. Then, the binarization of the ROI is performed using its average intensity as the threshold. Finally, the similarity is calculated by comparing the bit sequences. In this study, we defined arbitrary 30×30 square region extracted from the aligned images as the ROI and calculated the similarity.

2.2 Reconstruction Process

In the reconstruction process, we reconstruct an HR image from the observed LR images from the camera array system based on Farsiu's method [16]. According to Farsiu's model, the original HR image was degraded for several reasons and accordingly only the LR image was observed. This image degradation model is expressed as

$$Y_k = D_k H_k F_k X. \quad k = 1, \cdots, N. \tag{1}$$

Here, Y_k is the observed k-th LR image, D_k is the down-sampling process, H_k is the blurring function, F_k is the moving effect, X is the original (estimated) HR image, and N is the number of the input LR images.

We reconstructed the HR image based on the frequently used ML estimation method and a steepest descent method to solve the minimization problem of the differences between the original and estimated HR image. According to this, Eq. (1) can be rewritten as the following iterative formula.

$$X_{n+1} = X_n - \left\{ \sum_{k=1}^{N} F_k^T H_k^T D_k^T \mathrm{sign}\left(\mathrm{D}_k \mathrm{H}_k \mathrm{F}_k \mathrm{X}_n - \mathrm{Y}_k\right) \right\}. \tag{2}$$

Here, X_n is the n-th estimated HR image produced by steepest descent method. In the reconstruction process, we used the Butterworth low-pass filter as the blur kernel. Because of space limitation, for details, please refer the original article [16].

3 Super-Resolution Camera Array System

The prototype of our super-resolution camera-array system is shown in Fig. 2. This prototype is composed of 12 Axis M1011-W cameras. The resolution of each camera is 640×480 pixel. The horizontal and vertical viewing angle of each camera is 47° and 35°, respectively. The size of the overall system is 18 cm × 24 cm.

Since multiple cameras are arranged in our system, parallax and the differences in photographic conditions among the cameras cannot be ignored. We therefore propose three approaches to address these issues and improved the image quality.

Fig. 2. Camera array system.

Fig. 3. Observed image.

Fig. 4. Determined ROI image.

3.1 Color Calibration Among Cameras

We performed a color calibration among the observed images. We selected an arbitrary ROI from an HR image estimated as the basis for the calibration. We searched the corresponding areas among LR images in a manner similar to the procedure described in Sect. 2.1 and calibrated them using a histogram equalization strategy.

3.2 Automated ROI Detection-Strategy for Parallax

In a camera array system, parallax among equipped cameras occurs and it becomes particularly pronounced when the targets are located near the system. It causes the appearance of the objects to differ, in particular, in terms of depth. It is therefore quite important that the issue of parallax be treated appropriately. We propose a semi-automated detection method for the areas required for performing high resolution processing, i.e., the ROIs. First, we select the arbitrarily observed i_th image $(1, 2, \cdots, i, \cdots, N)$ and manually determine the area ROI_i that requires HR processing. Based on ROI_i, the corresponding areas among the observed images ROI_j $(1, 2, \cdots, j, \cdots, N, j \neq i)$ are automatically estimated using the image alignment procedure based on the method described in Sect. 2.1. An example of the observed image and the determined ROI_i is shown in Figs. 3 and 4, respectively.

3.3 Effective Noise Reduction with Effective Edge Preservation

Usually, a low-cost camera device comprises a small optical sensor (measuring "1/4", "1/2.33," similar). In situations where the illumination is insufficient, including even in normal indoor environments, gain noise is frequently incurred in the observed image produced by these sensor devices. Unfortunately, in some cases, this cannot be ignored in the super-resolution technique process. As

Fig. 5. Result of bi-linear method.

Fig. 6. Super-resolution result of conventional method from 12 LR images.

Fig. 7. Super-resolution result of proposed method from 12 LR images.

Fig. 8. Magnified view of result of bi-linear interpolation.

mentioned earlier, conventional super-resolution techniques frequently emphasize noise, as well as the requisite edges, contours, and so on, when the number of the observed images is limited. Since one of our objectives is to develop a low-cost super-resolution system, effective noise reduction is crucial.

At the beginning of our super-resolution process, the up-sampled observed LR image is treated as the initial estimation of the HR image. The initial HR image frequently includes noise originating from the base LR image. In addition, ML estimation during the iterative updating process sometimes amplifies noise, in particular when the given LR images have relatively large noise or the number of given LR images is limited. In order to address each of the above issues, we introduce the idea of using two filters; (1) a median filter applied to the initial HR image to reduce the initial noise, and (2) a bi-lateral filter applied to the sum of the residual between the HR image estimated and the i-th up-sampled LR image $(1, 2, \cdots, i, \cdots, N)$. It can be expected that introducing these two procedures will reduce noise, while preserving native edges, contours, and other important components. These filtering processes improved the image quality according to the qualitative visual assessment in our preliminary experiments.

Fig. 9. Magnified view of super-resolution result of conventional method from 12 LR images.

Fig. 10. Magnified view of super-resolution of proposed method from 12 LR images.

Fig. 11. Result of bi-linear method.

Fig. 12. Super-resolution result of proposed method from 12 LR images.

4 Results

First, we tested our system using a distant view. It should be noted that the magnification ratio in this experiment was 3×3. Figures 5, 6 and 7 show the results of bi-linear interpolation (Fig. 5), conventional super-resolution techniques (Fig. 6), and the proposed method (Fig. 7). Figures 8, 9 and 10 show enlarged areas of the images in Figs. 5, 6 and 7, respectively, for visual assessment.

Second, we tested our system using a short distance view. As explained previously, Fig. 4 illustrates the sample of the ROI in the LR image (ROIs were appropriately determined also in other LR images) and Figs. 11 and 12 show the results of bi-linear interpolation and the proposed system reconstructed from 12 LR images, respectively. In order to investigate the effectiveness of the proposed method, we compared the conventional super-resolution method and the proposed method based on only 4 LR images. The results are shown in Figs. 13 and 14, respectively. Figures 15, 16, 17 and 18 show enlarged areas of these images

Fig. 13. Super-resolution result of conventional method from 4 LR images.

Fig. 14. Super-resolution result of proposed method from 4 LR images.

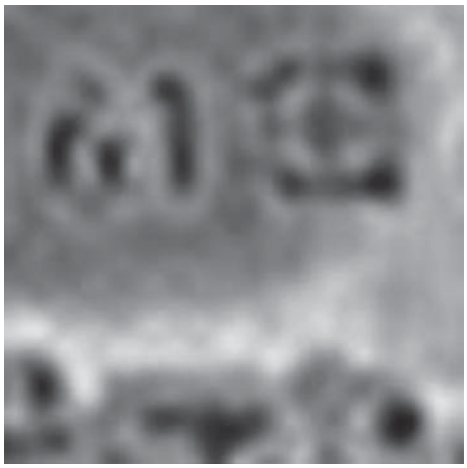

Fig. 15. Magnified view of bi-linear method.

Fig. 16. Magnified view of super-resolution result of conventional method from 12 LR images.

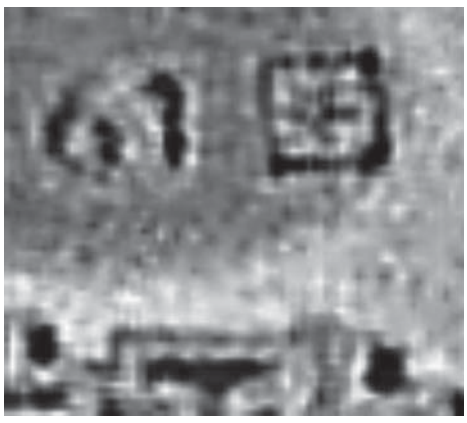

Fig. 17. Magnified view of super-resolution result pf conventional method from 4 LR images.

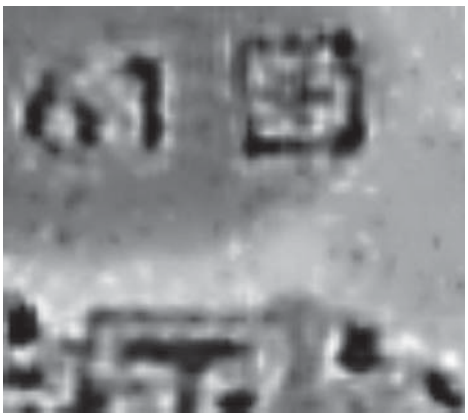

Fig. 18. Magnified view of super-resolution result of proposed method from 4 LR images.

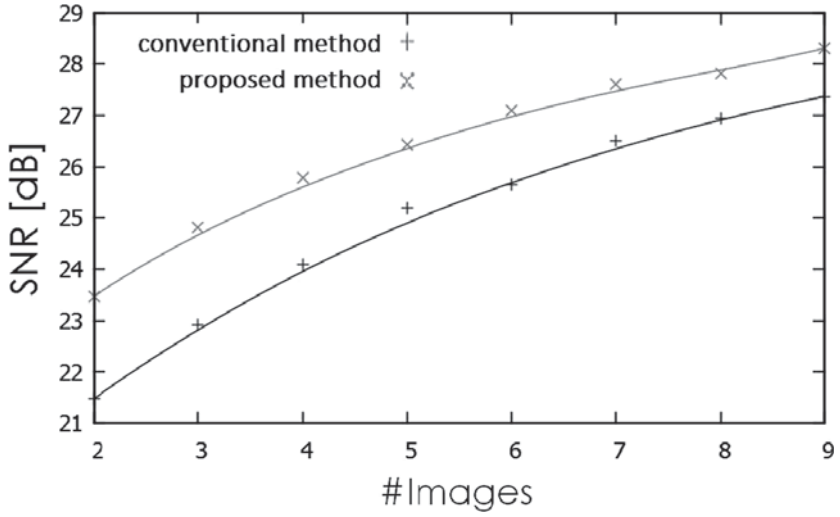

Fig. 19. Relationship between # number of images and the S/N ratio.

(Fig. 15: bi-linear, Fig. 16: proposed from 12 LR, Fig. 17: proposed method from 4 LR, and Fig. 18: conventional method from 4 LR).

As for the quantitative evaluation, we defined the result of the super-resolution using 40 LR images as the gold standard in this study and compared the proposed method and the conventional algorithm. Figure 19 illustrates the relationship between the number of base LR images and the S/N ratio. Note that, the S/N ratio of the bi-linear method was 20.1 dB.

It can be seen that the proposed method achieved superior results.

5 Discussion

We confirmed the advantages of the super-resolution technique in all cases, although in the cases of long distance imaging (see Figs. 6 and 7), the visual advantage of the proposed method over the conventional one was limited. We consider that the reasons are that the outdoor imaging with sufficient light did not produce a large gain noise in the original LR images and the number of the LR images used for reconstruction was sufficient.

In the short distance example, we confirmed that our proposed method successfully improved the resolution by resolving the parallax and color divergence issues (see Figs. 15 and 16). Now, we consider the cases where both the amount of the light and the number of reconstruction images were limited. It can be seen in Figs. 17 and 18 that the conventional method emphasizes the edges as well as noise, while the proposed method emphasizes the edges to the same degree as the conventional method, but successfully suppresses noise. In addition, it can be seen in Fig. 19 that the proposed method shows an S/N ratio superior to that of the conventional method by 2 dB. This is almost equivalent to the effect of two additional observed images.

The above results confirm that the proposed method is robust against noise and produces a high quality high-resolution image, even when the number of observed LR images is insufficient.

6 Summary

In this study, we developed a prototype of a super-resolution camera array system. The camera array system has several merits, on the other hand; however, on the other hand, it also suffers from several problem that need to be addressed to allow its practical usage. We proposed three processes to reduce the drawbacks of the array system and achieved an approximately 2 dB higher S/N ratio (i.e., equivalent to the effect of two additional images) than the conventional method.

The miniaturization of the system reduces the drawbacks of the array system. We will develop the next version of the model in the near future.

References

1. Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: a technical overview. IEEE Signal Process. Mag. **20**, 21–36 (2003)
2. Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-based super-resolution. IEEE Comput. Graphics Appl. **22**, 56–65 (2002)
3. Takashi, K., Takahiro, S.: Super-resolution decoding of the JPEG coded image data using total-variation regularization. In: Picture Coding Symposium, pp. 114–117 (2010)
4. Tom, B., Katsaggelos, A.: Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images. In: IEEE International Conference on Image Processing, vol. 2, pp. 539–542 (1995)
5. Schultz, R., Stevenson, R.: A bayesian approach to image expansion for improved definition. IEEE Trans. Image Process. **3**, 233–242 (1994)
6. Schultz, R., Stevenson, R.: Extraction of high-resolution frames from video sequences. IEEE Trans. Image Process. **5**, 996–1011 (1996)
7. Irani, M., Peleg, S.: Improving resolution by image registration. Graph. Models Image Process. **53**, 231–239 (1991)
8. Tsai, R., Huang, T.: Multiframe image restoration and registration. In: Advances in Computer Vision and Image, vol. 1, pp. 317–339 (1984)
9. Eren, P., Sezan, M., Tekalp, A.: Robust, object-based high resolution image reconstruction from low-resolution video. IEEE Trans. Image Process. **6**, 1446–1451 (1997)
10. Patti, A.J., Altunbasak, Y.: Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. IEEE Trans. Image Process. **10**, 179–186 (2001)
11. Zitova, B., Flusser, J.: Image registration methods: a survey. Image Vis. Comput. **21**, 977–1000 (2003)
12. Lowe, D.G.: Distinctive image features from scaleinva-invariant keypoints. Proc. Int. J. Comput. Vis. (IJCV) **60**, 91–110 (2004)
13. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (surf). Comput. Vis. Image Underst. **110**, 346–359 (2008)
14. Tung, T., Nobuhara, S., Matsuyama, T.: Simultaneous super-resolution and 3d video using graph-cuts. In: IEEE Conference Computer Vision Pattern Recognition, Anchorage (2008)
15. Aghav, S., Kumar, A., Gadakar, G., Mehta, A., Mhaisane, A.: Mitigation of rotational constraints in image based plagiarism detection using perceptual hash. Int. J. Comput. Sci. Trends Technol. **2**, 28–32 (2014)
16. Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multi-frame super-resolution. IEEE Trans. Image Process. **13**, 1327–1344 (2004)