Multiframe Super-Resolution for Flickering Objects

Atsushi Fukushima and Takahiro Okabe(✉)

Department of Artificial Intelligence, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
okabe@ai.kyutech.ac.jp

Abstract. In this paper, we propose a MAP-based multiframe super-resolution method for flickering objects such as LED electronic message boards. Since LED message boards often flicker at low refresh rates, missing areas where LEDs are off during the exposure time of a camera by chance are observed. To suppress unexpected artifacts due to those missing areas, our proposed method detects outlier pixels on the basis of the spatio-temporal analysis of pixel values, and removes them from the MAP estimation by incorporating the weights of pixels into the likelihood term. We conducted a number of experiments using both real and synthetic images, and qualitatively and quantitatively confirmed that our method works better than the existing methods.

Keywords: Image/video enhancement · Multiframe super-resolution · Maximum a posteriori estimation · Outlier removal · Electronic message board

1 Introduction

Electronic message boards are used for displaying various important and/or useful information such as news, weather forecasts, road traffic information and signs, arrival and departure information of flights and trains, and advertisements. Along with the popularization of mobile phones with cameras and car-mounted cameras, we have greater opportunities to capture the image sequences of electronic message boards. Since those image sequences are often low resolution, improving their image quality makes it easier for both us and computers to understand the messages displayed on them.

Most electronic message boards consist of LED arrays. In a similar manner to CRT displays, LED message boards often flicker at low refresh rates. When we capture an image sequence of such a flickering object with a long exposure time, flickering is not observed, but the high-frequency components of the images are lost due to motion blur. Recovering the high-frequency components of blurred images is still a challenging problem to be addressed. On the other hand, when capturing with a short exposure time, motion blur is reduced, but flickering is observed. For example, band-like dark areas, *i.e.* missing areas are observed in

G. Azzopardi and N. Petkov (Eds.): CAIP 2015, Part II, LNCS 9257, pp. 184–194, 2015.
DOI: 10.1007/978-3-319-23117-4_16

Fig. 1. Real images of an electronic message board: band-like dark areas are observed due to flickering.

Fig. 1; the LEDs in those areas are off during the exposure time of a camera by chance. In this paper, we address a novel super-resolution problem; super-resolution for flickering objects when missing areas are observed.

In general, super-resolution techniques are classified into two approaches; a learning-based one and a reconstruction-based one. The learning-based approach makes use of a set of example images, and is able to enhance image quality even from a single input image. However, when missing areas are observed in the single image, it is difficult, if not impossible, to detect and restore missing areas. Therefore, we address a reconstruction-based approach, which makes use of multiple input images of the same object, to super-resolution for flickering objects. In particular, we consider multiframe super-resolution based on MAP (Maximum-A-Posteriori) estimation[4].

Conventionally, multiframe super-resolution assumes that the brightness of an object of interest is time-invariant. However, the brightness of a flickering object such as an electronic message board changes temporally, and as a result the performance of the conventional multiframe super-resolution is significantly degraded when missing areas are observed. Accordingly, our proposed method detects those missing areas, *i.e.* outlier pixels on the basis of the spatio-temporal analysis of pixel values, and then removes those outlier pixels from the MAP estimation by incorporating the weights of pixels into the likelihood term. We conducted a number of experiments by using both real and synthetic images, and qualitatively and quantitatively confirmed that the proposed method works well for flickering objects.

Robust super-resolution is one of the most important research topics in image/video enhancement[1,2,7,8]. It has already been reported that outlier removal is effective for suppressing unexpected artifacts due to registration errors[7] and video compression[2]. The main contribution of this study is twofold; we experimentally demonstrate that (i) the framework of outlier removal is effective also for suppressing unexpected artifacts due to flickering objects such as electronic message boards, and that (ii) the spatio-temporal analysis of pixel values works well for detecting outliers caused by flickering.

The rest of this paper is organized as follows. We briefly summarize related work in Section 2. A MAP-based multiframe super-resolution method for flickering objects such as LED electronic message boards is proposed in Section 3. We report the experimental results in Section 4 and present concluding remarks in Section 5.

2 Related Work

Existing methods for robust super-resolution can be classified into two categories; one is based on robust estimator and the other is based on outlier removal. In this section, we briefly explain the existing methods in each category, and then describe the relationship to our proposed method.

The former approach makes use of robust estimator, which is insensitive to outliers. In order to reduce the effects of noises and errors in motion model and blur model, Farsiu *et al.*[1] propose a robust super-resolution method by using the L1-norm likelihood term and the edge-preserving regularization term. Zomet *et al.*[8] incorporate median estimator into the optimization of super-resolution reconstruction. Those methods can reduce unexpected artifacts due to outlier pixels without explicitly detecting them. However, it is difficult to remove the effects of outlier pixels throughly.

The latter approach explicitly detects outlier pixels and removes them from super-resolution reconstruction. Zhao and Sawhney[7] propose an optical flow based super-resolution method, and make use of outlier removal in order to get rid of registration errors. The framework of outlier removal is used also for dealing with multiple motions in a scene[6]. Ivanovski *et al.*[2] propose a super-resolution method robust to noises due to video compression.

Ivanovski's method is most closely related to ours, because it detects outliers pixel-wisely on the basis of the temporal analysis of pixel values. Specifically, they compute the median of pixel values at corresponding pixels in the observed image sequence, and then consider a pixel in a certain frame as outlier if the difference between its pixel value and the median is larger than a threshold. Unfortunately, however, their method is not suited for flickering objects, because the median can be outlier when missing areas are observed in more than half of the frames. In addition, we need to tune the threshold for detecting outliers manually.

3 Proposed Method

In this section, we propose multiframe super-resolution for flickering objects such as electronic message boards when missing areas are observed. We describe how the framework of outlier removal is used for dealing with missing areas due to flickering, and then describe how to detect outliers on the basis of the spatio-temporal analysis of pixel values.

3.1 Multiframe Super-Resolution with Outlier Removal

The relationship between a desired high-resolution image and observed low-resolution images is described by the observation model;

$$\boldsymbol{y}_f = \boldsymbol{D}\boldsymbol{B}_f\boldsymbol{M}_f\boldsymbol{x} + \boldsymbol{n}_f. \tag{1}$$

Fig. 2. Pixels in an observed low-resolution image are classified into three categories; true foreground, false background, and true background.

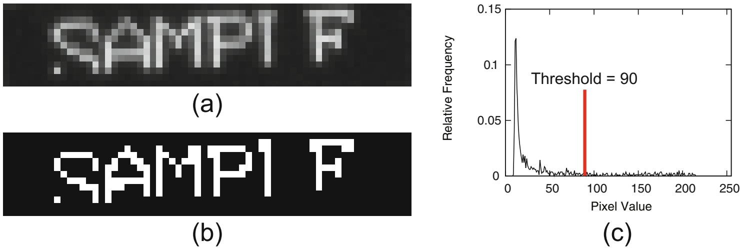

Fig. 3. True foreground detection based on spatial analysis: an observed image (a) is binarized by thresholding (b). The threshold is automatically determined from the histogram of pixel values (c) in the observed image on the basis of discriminant criterion. White pixels stand for true foreground.

Here, $\boldsymbol{x}$ and $\boldsymbol{y}_f$ $(f = 1, 2, 3, ..., F)$ are the vectors representing the desired high-resolution image and observed low-resolution images, and F is the number of frames, *i.e.* the number of the observed images. $\boldsymbol{M}_f$, $\boldsymbol{B}_f$, and $\boldsymbol{D}$ are matrices representing the (camera) motion of the f-th frame, the point spread function (PSF) of the f-th frame, and the down sampling respectively. $\boldsymbol{n}_f$ is a vector representing noises in the f-th observed image. Hereafter, we describe the observation model as

$$\boldsymbol{y}_f = \boldsymbol{A}_f \boldsymbol{x} + \boldsymbol{n}_f, \tag{2}$$

where $\boldsymbol{A}_f = \boldsymbol{D}\boldsymbol{B}_f\boldsymbol{M}_f$ for the sake of simplicity.

Assuming that the noises obey the zero-mean Gaussian distribution, the cost function for multiframe super-resolution based on MAP estimation is given by

$$c(\boldsymbol{x}) = \sum_{f=1}^{F} ||\boldsymbol{y}_f - \boldsymbol{A}_f \boldsymbol{x}||_2^2 + \lambda ||\boldsymbol{H}\boldsymbol{x}||_2^2, \tag{3}$$

and the high-resolution image is estimated by minimizing the cost function $c(\boldsymbol{x})$ with respect to $\boldsymbol{x}$. Here, the first and second terms represent the likelihood and prior knowledge of the high-resolution image, and λ is the parameter that balances those two terms. In order to obtain a smooth high-resolution image, high-pass filters are often used for $\boldsymbol{H}$.

We assume that an object of interest is static except for flickering. Specifically, our proposed method assumes that characters and/or symbols themselves displayed on an electronic message board are fixed during capturing an image

Fig. 4. Classification of true background and false background based on temporal analysis: we consider a dark pixel in a certain frame as (a) true background if all of the corresponding pixels in the different frames are also dark or (b) false background if at least one of the corresponding pixel in the different frame is bright.

sequence[1]. The objective of our method is to improve the image quality of those characters and symbols, *i.e.* to estimate the high-resolution image $\boldsymbol{x}$ without missing areas.

As shown in Fig. 2, the pixels in the observed low-resolution image are classified into three categories; *true foreground*, *false background*, and *true background*. The true foreground pixels correspond to the LEDs composing characters and symbols, and those LEDs are on during the exposure time of a camera. The false background pixels also correspond to the LEDs composing characters and symbols, but those LEDs are off during the exposure time by chance. The true background pixels correspond to the remaining LEDs, and those LEDs are always off. The true foreground pixels and true background pixels are considered to satisfy the observation model with the zero-mean Gaussian noises in eq.(2). On the other hand, the false background pixels significantly deviate from the observation model.

Accordingly, our proposed method makes use of the framework of outlier removal[2, 7]. Specifically, the proposed method removes outlier pixels from the MAP estimation by incorporating the weights of pixels into the likelihood term as

[1] If necessary, segmentation via image subtraction could be used for dynamic messages.

$$c'(\boldsymbol{x}) = \sum_{f=1}^{F} (\boldsymbol{y}_f - \boldsymbol{A}_f \boldsymbol{x})^\top \mathrm{diag}(\boldsymbol{w}_f)(\boldsymbol{y}_f - \boldsymbol{A}_f \boldsymbol{x}) + \lambda ||\boldsymbol{H}\boldsymbol{x}||_2^2. \tag{4}$$

Here, $\mathrm{diag}(\boldsymbol{w}_f)$ is the diagonal matrix whose diagonal elements are the elements of $\boldsymbol{w}_f$. If the p-th pixel ($p = 1, 2, 3, ..., P$) in the f-th frame is an inlier (outlier), $w_{f,p} = 1$ ($w_{f,p} = 0$). Our method considers the true foreground pixels and true background pixels as inliers, and considers the false background pixels as outliers, and then estimates the high-resolution image by minimizing the cost function $c'(\boldsymbol{x})$ with respect to $\boldsymbol{x}$.

3.2 Outlier Detection Based on Spatio-Temporal Analysis

We assume that a region of interest (ROI), *i.e.* the region corresponding to an electronic message board is cropped in advance. Then, we detect outliers on the basis of the spatio-temporal analysis of pixel values. In other words, we classify the pixels in the ROI into true foreground, false background, and true background.

First, we detect true foreground pixels on the basis of the spatial analysis of pixel values. Specifically, we consider bright pixels, whose pixel values are larger than a certain threshold, as true foreground as shown in Fig. 3. According to Otsu's method [3], the threshold is automatically determined from the histogram of pixel values in the ROI of each observed image on the basis of discriminant criterion. Note that dark pixels, whose pixel values are smaller than the threshold, can be either false background or true background.

Second, we classify the remaining pixels into false background and true background on the basis of the temporal analysis of pixel values. Specifically, we consider a dark pixel in a certain frame as true background, if all of the corresponding pixels in the different frames are also dark as shown in Fig. 4 (a). On the other hand, we consider a dark pixel in a certain frame as false background, if at least one of the corresponding pixel in the different frame is bright as shown in Fig. 4 (b).

4 Experiments

To demonstrate the effectiveness of our proposed method, we conducted a number of experiments using both real and synthetic images. We compared the performance of our method with those of existing methods; Pickup *et al.*[5] and Ivanovski *et al.*[2].

We used Pickup *et al.*[5] as a conventional method without outlier removal, and used Ivanovski *et al.*[2] as a robust method with outlier removal because it is the most closely related work to ours. Ivanovski *et al.* detect outlier pixels on the basis of the temporal analysis of pixel values. Specifically, they compute the median of pixel values at corresponding pixels in the observed image sequence, and then consider a pixel in a certain frame as outlier if the difference between

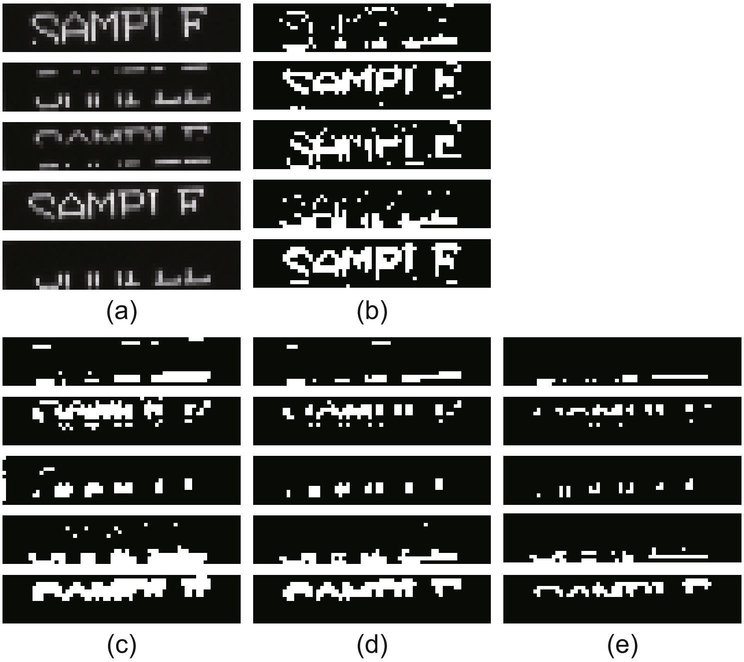

Fig. 5. The observed low-resolution images (a), the weights of pixels given by our proposed method based on the spatio-temporal analysis and automatic thresholding (b), and those given by the existing method based on the median and manual thresholding: (c) $t = 0.05$, (d) $t = 0.1$, and (e) $t = 0.2$. White pixels stand for outliers with $w = 0$.

its pixel value and the median is larger than a threshold t. Here, the threshold for detecting outliers is empirically determined.

Both our proposed method and Ivanovski *et al.*[2] were implemented by incorporating the outlier detection and removal into the implementation of Pickup *et al.*[5] for fair comparison. Note that they could be combined with other existing methods for multiframe super-resolution based on MAP estimation.

4.1 Experiments Using Real Images

We conducted experiments using the real images of two different LED electronic message boards. Those images were captured by using a Point Grey Chameleon camera with a global shutter. The exposure time was 5.3[ms]. The number of frames in each image sequence was 5 (P=5). We assumed planar homography estimated by using markers and achieved the registration of those images[2].

In Fig. 5, we show the observed low-resolution images (a), the weights of pixels given by our proposed method based on the spatio-temporal analysis and

[2] In natural scenes, time-invariant feature points such as the corners of an electronic message board could be used for registration.

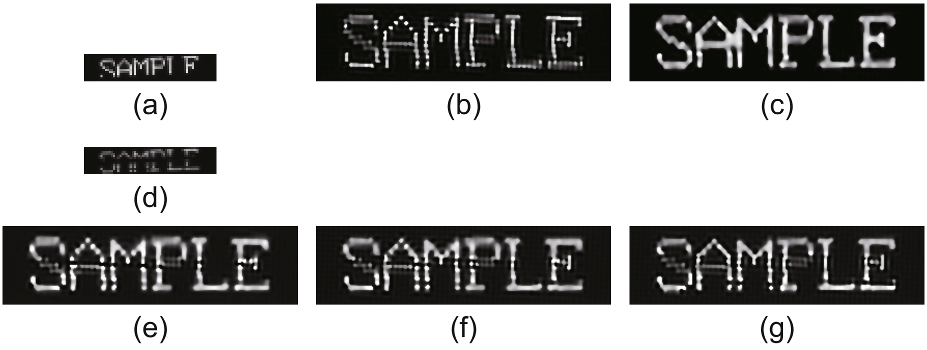

Fig. 6. Super-resolution reconstruction (SAMPLE): an input image (a), the conventional method without outlier removal (b), our proposed method (c), the median image (d), and the median-based method with (e) $t = 0.05$, (f) $t = 0.1$, and (g) $t = 0.2$.

automatic thresholding (b), and those given by the existing method based on the median and manual thresholding: (c) $t = 0.05$, (d) $t = 0.1$, and (e) $t = 0.2$[3]. Here, white pixels stand for outliers with $w = 0$. We can see that our method works well although slightly excessive amount of outliers are detected. On the other hand, we can see that the amount of outliers detected by the median-based method significantly depends on the empirically-determined threshold. More importantly, some of false background pixels are misclassified into inliers, because the median image is wrong, *i.e.* missing areas are observed in the median image shown in Fig. 6 (d).

In Fig. 6, we show an input image from the observed image sequence (a), the super-resolution reconstruction by using the conventional method without outlier removal (b), and our proposed method (c). In addition, we show the median image (d), and the super-resolution reconstruction by using the median-based method with (e) $t = 0.05$, (f) $t = 0.1$, and (g) $t = 0.2$. We can see that our method works better than the conventional method without outlier removal. Specifically, the conventional method tries to explain the temporal brightness variations due to flickering as the degradation of high-frequency grained patterns, and as a result annoying artifacts due to flickering are visible. We can also see that the median-based method does not work well, in particular where missing areas are observed in the median image. In addition, the performance of the median-based method depends on the empirically-determined threshold.

Fig. 7 and Fig. 8 show the results for other image sequences; different LED message boards, different characters and symbols, and different imaging conditions. We can see that our proposed method outperforms the conventional method without outlier removal[5] and the median-based method with an empirically-determined threshold[2]. Therefore, we can conclude that those experimental results qualitatively demonstrate the effectiveness of our method.

[3] 8-bit pixel values are normalized to $[0, 1]$.

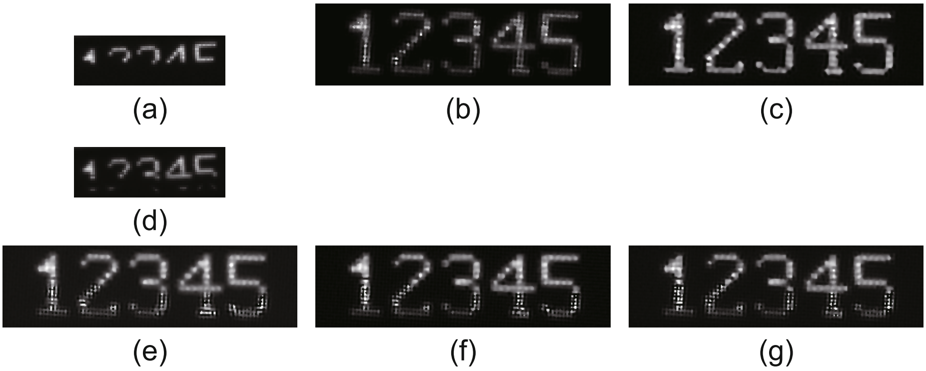

Fig. 7. Super-resolution reconstruction (12345): an input image (a), the conventional method without outlier removal (b), our proposed method (c), the median image (d), and the median-based method with (e) $t = 0.05$, (f) $t = 0.1$, and (g) $t = 0.2$.

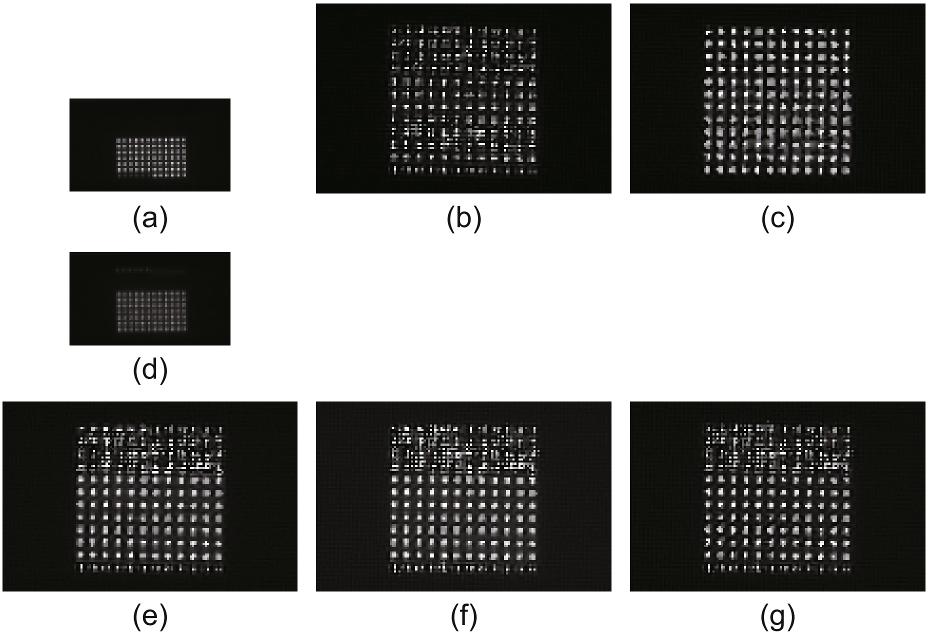

Fig. 8. Super-resolution reconstruction (square): an input image (a), the conventional method without outlier removal (b), our proposed method (c), the median image (d), and the median-based method with (e) $t = 0.05$, (f) $t = 0.1$, and (g) $t = 0.2$.

4.2 Experiments Using Synthetic Images

In order to quantitatively demonstrate the effectiveness of our proposed method, we conducted experiments using synthetic images under the condition that the

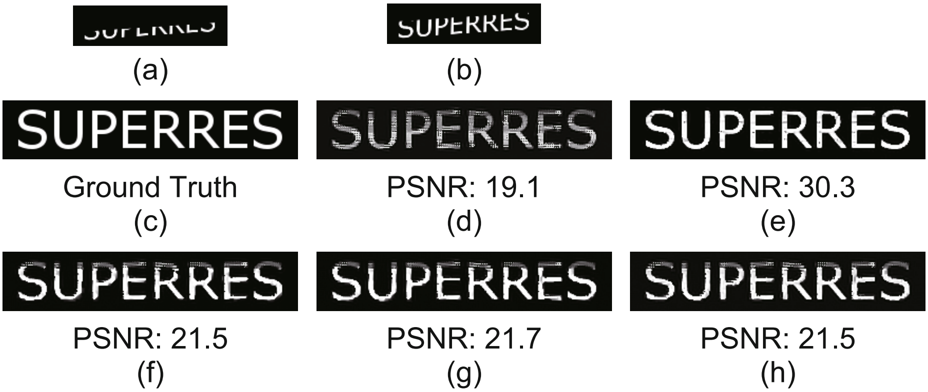

Fig. 9. Super-resolution reconstruction (synthetic): an input image (a), the median image (b), the ground truth of a high-resolution image (c), the conventional method without outlier removal (d), our proposed method (e), and the median-based method with (f) $t = 0.05$, (g) $t = 0.1$, and (h) $t = 0.2$.

ground truth of a high-resolution image is known. The number of frames was 5 (P=5). We assumed that planar homography is also known.

In Fig. 9, we show an input image from the synthetic image sequence (a), the median image (b), the ground truth of a high-resolution image (c), the super-resolution reconstruction by using the conventional method without outlier removal (d), our proposed method (e), and the median-based method with (f) $t = 0.05$, (g) $t = 0.1$, and (h) $t = 0.2$. We also show the PSNR (Peak Signal-to-Noise Ratio) of each reconstructed image. Similar to the results on the real images, we can see that our method works better than the conventional method without outlier removal[5] and the median-based method with an empirically-determined threshold[2]. Those results quantitatively demonstrate the effectiveness of our method.

5 Conclusion and Future Work

In this paper, we proposed a multiframe super-resolution method for flickering objects such as electronic message boards. Specifically, our proposed method detects outlier pixels due to flickering on the basis of the spatio-temporal analysis of pixel values, and then removes them from the MAP estimation by incorporating the weights of pixels into the likelihood term. We conducted a number of experiments using both real and synthetic images, and demonstrated that our method works better than the existing methods. The future work of this study includes the extension to dynamic messages and the extension to image sequences in which not only flickering but also motion blur are observed.

Acknowledgments. We would like to thank Prof. Tsukasa Noma for his valuable comments and suggestions early in this project.

References

1. Farsiu, S., Robinson, M., Elad, M., Milanfar, P.: Fast and robust multiframe super resolution. IEEE Trans. Image Processing **13**(10), 1327–1344 (2004)
2. Ivanovski, Z., Panovski, L., Karam, L.: Robust super-resolution based on pixel-level selectivity. In: Proc. SPIE 6077, Visual Communications and Image Processing 2006, 607707 (2006)
3. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Systems, Man, and Cybernetics **SMC–9**(1), 62–66 (1979)
4. Park, S., Park, M., Kang, M.: Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine **20**(3), 21–36 (2003)
5. Pickup, L., Roberts, S., Zisserman, A.: A sampled texture prior for image super-resolution. In: Proc. NIPS 2003, pp. 1587–1594 (2003)
6. Tanaka, M., Yaguchi, Y., Okutomi, M.: Robust and accurate estimation of multiple motions for whole-image super-resolution. In: Proc. IEEE ICIP 2008, pp. 649–652 (2008)
7. Zhao, W., Sawhney, H.S.: Is super-resolution with optical flow feasible? In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part I. LNCS, vol. 2350, pp. 599–613. Springer, Heidelberg (2002)
8. Zomet, A., Rav-Acha, A., Peleg, S.: Robust super-resolution. In: Proc. IEEE CVPR 2001, pp. I-640–I-650 (2001)