# Removal of Adherent Waterdrops from Images Acquired with Stereo Camera*

Atsushi Yamashita, Yuu Tanaka and Toru Kaneko
*Department of Mechanical Engineering*
*Shizuoka University*
*3–5–1 Johoku, Hamamatsu-shi, Shizuoka 432–8561, Japan*
*yamashita@ieee.org*

*Abstract*— In this paper, we propose a new method that can remove view-disturbing noises from stereo images. One of the thorny problems in outdoor surveillance by a camera is that adherent noises such as waterdrops on the protecting glass surface lens disturb the view from the camera. Therefore, we propose a method for removing adherent noises from stereo images taken with a stereo camera system. Our method is based on the stereo measurement and utilizes disparities between stereo image pair. Positions of noises in images can be detected by comparing disparities measured from stereo images with the distance between the stereo camera system and the glass surface. True disparities of image regions hidden by noises can be estimated from the property that disparities are generally similar with those around noises. Finally, we can remove noises from images by replacing the above regions with textures of corresponding image regions obtained by the disparity referring. Experimental results show the effectiveness of the proposed method.

*Index Terms*— Image restoration, Stereo images, Noise removal, Template matching, Disparity estimation

## I. INTRODUCTION

In recent years, surveillance systems using cameras are widely used for the traffic flow observation, the trespassers detection, and so on, owing to the performance improvement and the cost reduction in computers and image input devices. The task that mobile robots collect the information about the environment by using a camera also will become very significant and be in high demand for security or disaster response in the near future. In these cases, automatic surveillance and recognition systems are expected because it is very difficult for human operators to check the situation at all times.

However, the qualities of images taken through cameras depend on environmental conditions. It is often the case that scenes taken by the cameras in outdoor environments are difficult to see because of adherent noises on the surface of the lens-protecting glass of the camera. For example, waterdrops attached on the protecting glass may interrupt a field of view (FOV) in rainy days. It would be desirable to remove adherent noises from images of such scenes for the surveillance and the environment recognition. Especially in a disaster, it is one
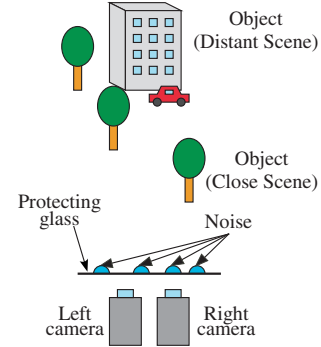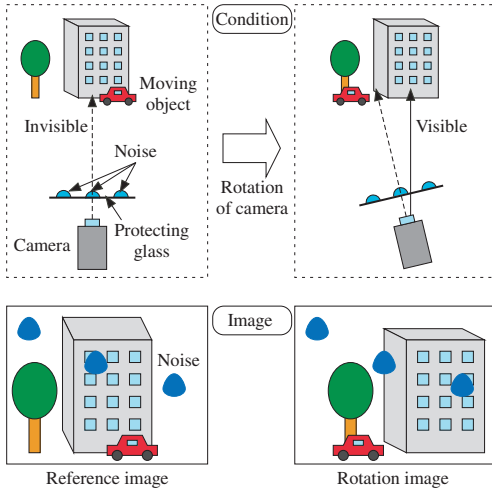
Fig. 1. Overview of our method.

of the most important things for rescue robots to maintain a clear view for the rapid and reliable search.

Therefore, this paper proposes a new method for removing adherent noises from images acquired with a stereo camera system (Fig. 1).

The detection of noise positions in images and the interpolation of these adherent areas are essential techniques to solve this problem.

As to the detection of the position of noise areas in images, there are a lot of studies that detect moving objects or noises in images [1]–[5]. These techniques remove moving objects or noises by taking the difference between the initial background scene and a current scene (background subtraction), or taking the difference between temporarily adjacent two frames (interframe subtraction). These methods are robust against the change of background [2], the change of the weather [3], or the change of the lighting condition [4]. An efficient algorithm for detecting and removing rain from videos based on a physics-based motion blur model that explains the photometry of rain is also proposed [5]. However, the methods based on the background subtraction have a disadvantage that it cannot be used in case when the background itself changes. The methods based on the interframe subtraction also have a disadvantage that it cannot detect stationary objects after they have appeared and stay in the image. Therefore, it is difficult to apply these techniques to the above problem, because they cannot detect stationary noises and adherent noises such as waterdrops may be stationary noises in the images.

(a) Overview.



(b) Original image.          (c) Improved image.

Fig. 2.   Restoration of deteriorated images by using a pan-tilt cameras [13].



(a) Overview.
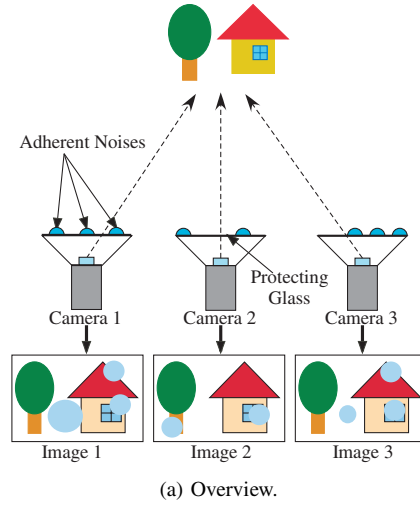


(b) Original image.          (c) Improved image.

Fig. 3.   Restoration of deteriorated images by using multiple cameras [15].

On the other hands, a lot of image interpolation or restoration techniques for damaged and occluded images are also proposed [6]–[11]. However, some of them can only treat with line-shape scratches [6]–[8], because they are the techniques for restoring old damaged films. It is also required that human operators indicate the region of noises interactively (not automatically) [9]–[11]. At any hand, it is also very difficult to treat large noises and to duplicate the complex textures with these methods.

To solve these problems, we have proposed the method that can remove view-disturbing noises from images taken with a pan-tilt camera [12]–[14]. This method is based on the comparison of two images, a reference image and a second image taken by a different camera angle (Fig. 2). However, it assumes that waterdrops never change their positions in images while the camera rotates. Therefore, it is not strong against the situations of heavy rain days.

We have also proposed the noise removal method by using more than two cameras [15]. This method is based on the comparison of images that are taken with multiple cameras, and there is no assumption about waterdrop movement (Fig. 3). However, it cannot be used for close scenes that have disparities between different viewpoints, because it is based on the difference between images.

Stereo camera systems are widely used for robot sensors, and they must of course observe both distant scenes and close scenes. Therefore, this paper proposes a new method that can remove waterdrops from stereo image pairs that contain objects both in a distant scene and in a close range scene (Fig. 1). Our new method can clear up above-mentioned problem of the previous method [15].

The composition of this paper is detailed below. In Section II, we mention about outline of our method. In Section III, the constraints of the configuration among the protecting glass and the cameras are considered. In Section IV, the method of detecting noise positions is explained, and in Section V, the image correction method is constructed. In Section VI, experimental results are shown and we discuss the effectiveness of our method. Finally, Section VII describes conclusions and future works.

## II. OUTLINE OF OUR METHOD

Our proposed method consists of the following three steps.

1) Stereo image pairs are acquired with a stereo camera system. They are transformed to parallel stereo image pairs by positional and chromatic registrations [13].
2) Template matching with the normalized cross correlation is performed between images of a gray-scale stereo pair. Noises are distinguished by using disparity and correlation of each pixel.
3) Disparities of areas where disparities are not given by the matching process are interpolated. Noises existing in a common FOV of a stereo image pair are removed by replacing its pixels with the corresponding textures in the other image obtained by referring their disparities.
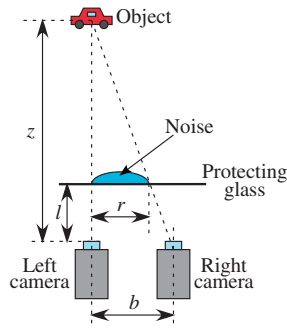
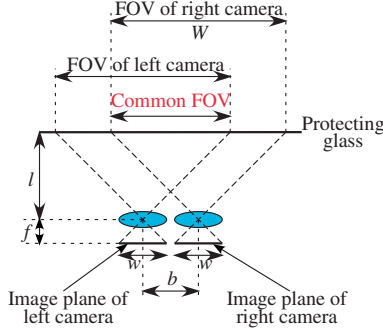Fig. 4.   Constraints of camera configuration about noise removal.



Fig. 5.   Constraints of camera configuration about common FOV.

## III. CONSTRAINTS OF CAMERA CONFIGURATION

The proposed method removes noises in the common FOV of a stereo image pair by replacing pixels of noises in one image with pixels in the other image. Cameras view direction is same and a protection glass surface is set up in perpendicular in the cameras view direction. Optical geometry of cameras is shown in Fig. 4, since the cameras view direction is same in the case of parallel stereo images. The baseline length needs to satisfy Eq.(1), because a background object must be observed with at least one camera.

$$b > \frac{zr}{z-l}, \tag{1}$$

where $b$, $l$, $z$ and $r$ denote the baseline length, the distance between cameras and a protection glass, the distance between the cameras and an object nearest to the cameras, and a noise size, respectively.

The rate of the common FOV to the original FOV of each camera must be larger than certain threshold $E$ (Fig. 5). This constraint can be expressed as follows:

$$\frac{W - b}{W} > E, \tag{2}$$

where $W$ is the original FOV of each camera.

Equation (2) can be transformed to Eq.(3) because $W = lw/f$.

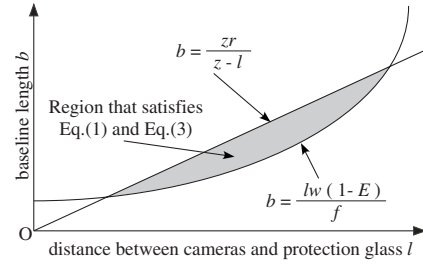$$1 - \frac{f}{lw}b > E, \tag{3}$$



Fig. 6.   Constraints of camera configuration.

where $f$ is an image distance[1] and $w$ is the image plane size, respectively.

From Eq.(1) and Eq.(3), we can obtain the constraints of the configuration among the protecting glass and two cameras (Fig. 6). The baseline length $b$ and the distance between the cameras and the protection glass $l$ must satisfy these constraints.

## IV. DETECTION OF NOISE POSITION

### A. Template Matching

The proposed method performs template matching by normalized cross correlation (NCC) of the stereo images. Correlation value $R$ of NCC is calculated as follow:

$$R = \frac{\sum_{j=1}^{N} \sum_{i=1}^{M} (I_l(i,j) - \mu_l)(I_r(i,j) - \mu_r)}{M N \sigma_l \sigma_r}, \tag{4}$$

where $I_{r,l}(i,j)$ is the pixel value of the left and right gray-scale image at pixel $(i,j)$, $\mu_{l,r}$ and $\sigma_{l,r}$ are average and standard deviation of pixel value of templates, and $M \times N$ is a template size, respectively.

$$\mu_{l,r} = \frac{1}{NM} \sum_{j=1}^{N} \sum_{i=1}^{M} I_{l,r}(i,j), \tag{5}$$

$$\sigma_{l,r}^2 = \frac{1}{NM} \sum_{j=1}^{N} \sum_{i=1}^{M} (I_{l,r}(i,j) - \mu_{l,r})^2. \tag{6}$$

Then disparities and correlations are acquired, and noise positions are estimated.

### B. Detection by One-to-One Correspondence

The positions of noises are detected by using disparities that are obtained by template matching of stereo images. Template matching causes errors, when intensity variation in a template is little, or when a matching point does not exist by occlusion. Therefore, it is necessary to investigate a reliability of template matching results. In order to investigate the reliability, two criteria are adopted here.

---

[1]The image distance is equal to the distance between the center of lens and the image plane. Although it is confusable, the image distance is not same as the focal length. When an image of an infinitely (or at least sufficiently) distant object is created on the sensor, this distance is equal to the focal length of the lens [16].

One is thresholding of correlations. If a correlation $R$ is less than a threshold $C$, the matching result is discarded as unreliable.

The other criterion is investigating whether results of template matching correspond one-on-one. If a matching result is correct, it corresponds one-on-one. Suppose that a pixel at $(u, v)$ in one image is set as a center of a template, and a matched pixel is found at $(u', v')$ in the second image. Next, template matching is again performed by setting a pixel at $(u', v')$ in the second image as a center of a template. The result has one-to-one correspondence only when $(u, v) = (u'', v'')$, where $(u'', v'')$ is the coordinates of the matched pixel in the first image. However, we should give some tolerance for this condition because of an image noise. Pixel $(u, v)$ is given a judgment value $\gamma(u, v)$ by Eq.(7).

$$\gamma(u, v) = \begin{cases} 1, & R \geq C \ and \ |u - u''| + |v - v''| \leq \xi \\ 0, & otherwise \end{cases}, \quad (7)$$

where $\xi$ is a threshold value.

When $\gamma(u, v)$ is 1, a similarity of a matching result is high, and the result corresponds uniquely.

### C. Noise Distinction by Disparity

Noises adhere on a protection glass surface. Therefore, disparity of the noises can be calculated from camera parameters and geometrical relation between the protection glass and the stereo camera system.

Disparity of the noise region $\eta$ is calculated from Eq.(8).

$$\eta = \frac{bf}{l}. \quad (8)$$

Disparity $S(u, v)$ is calculated from a matching result when $\gamma(u, v)$ is 1, and $S(u, v)$ is compared to disparity $\eta$. We set a threshold $\delta$ for distinguishing whether it is noise or not. Pixels of $|S(u, v) - \eta| < \delta$ are regarded as noise elements.

In Eq.(9), $\alpha(u, v)$ is the result of noise detection given to each pixel. Pixels of $\alpha(u, v) = 1$ are noise elements.

$$\alpha(u, v) = \begin{cases} 1, & \gamma(u, v) = 1 \ and \ |S(u, v) - \eta| < \delta \\ 0, & otherwise \end{cases}. \quad (9)$$

### V. Image Correction

Hidden regions by noises on the protecting glass are usually given in the other image. Therefore, noises can be removed by replacing the pixel intensities with those in the other image. In order to use the pixel intensities of the other image for a noise removal, the positions corresponding to the noises are required. Therefore, it is necessary to estimate disparities in positions of noises.

### A. Disparity Estimation

Disparities in noise positions are estimated with the "image inpainting" algorithm [9]. Originally, this method corrects the noise of an image in consideration of slopes of image intensities. The merit of this algorithm is the fine reproducibility for edges and its demerit is the poor reproducibility for a complicated texture. The proposed method in this paper

treats a disparity $S(u, v)$ as a pixel intensity and estimates disparities of pixels of $\alpha(u, v) = 1$ by using the image inpainting algorithm. In many cases disparities do not produce a complicated texture than intensities. Therefore, the demerit of poor reproducibility for complicated texture can be ignored.

### B. Image Interpolation

Noises are removed after the estimation of disparities. A pixel intensity $I(u, v)$ in a noise position is given by the following equation, where $s(u, v)$ is the estimated disparity and $I'(u, v)$ is the pixel intensity of the complementary image in the image pair.

$$I(u, v) = \begin{cases} I'(u - s, v), & (u, v) \ is \ in \ left \ image \\ I'(u + s, v), & (u, v) \ is \ in \ right \ image \end{cases}. \quad (10)$$

### VI. Experiment

We verified the effectiveness of the proposed method through experiments. The resolutions of all images were set as $640 \times 480$pixels.

Figure 7(a)(b) shows the original stereo images of a scene that consists of objects with a variety of distance when there are waterdrops on the protecting glass. Figure 7(c)(d) shows positions of waterdrops indicated manually for reference.

In this experiment, the image distance $f$ equals to 715pixel and the image plane size $w$ equals to 640pixel. The conditions of the image acquisition was set as follows: minimum object distance $z = 450$mm, common FOV rate $E = 0.85$. We also set 10mm as the largest noise size $r$, because waterdrop whose radius is more than 10mm always runs down. The constraints of the camera configuration can be calculated by substituting these parameters in Eq.(1) and Eq.(3). We decided the camera configuration that the distance between the protection glass and the cameras $l$ was 210mm and the baseline length $b$ was 20mm under the condition that the above-mentioned constraints were fulfilled.

The disparity $\eta$ for the protection glass surface calculated using Eq.(8) was 79pixel. The template size in template matching was $11 \times 11$pixels. The threshold $C$ for a correlation value was 0.4. The threshold $\xi$ that investigates one-to-one correspondence was 4. The threshold $\delta$ for noise detection was 10.

Figure 8 shows the results of waterdrop position detection and the disparity estimation. White areas in Fig. 8(a)(b) indicates waterdrop regions. Red pixels have large disparities (close scene) and green pixels have small disparities (distant scene). Black pixels have unknown disparities. Estimation results of disparities of while and black pixels are shown in Fig. 8(c)(d). Figure 9 shows results of waterdrop removal, and Fig. 10 shows magnified left images of waterdrop removal results. It is possible to read Chinese characters in the improved images, although it is impossible to read them in the original images.

Figure 11 shows the result that contains a very close object. From this result, it is verified that our method can work well in the case of very close scenes.

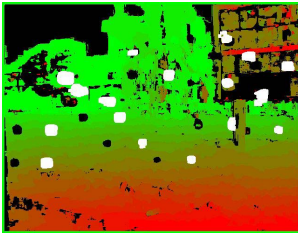(a) Left image.

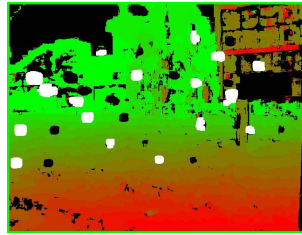(b) Right image.

(c) Noise position in (a).
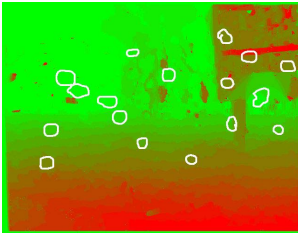
(d) Noise positions in (b).
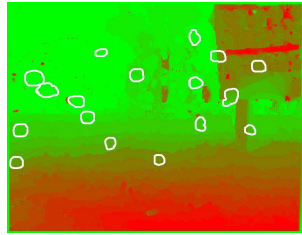
Fig. 7. Original stereo images.



(a) Disparity of left image.

(b) Disparity of right image.

(c) Disparity estimation of (a).

(d) Disparity estimation of (b).

Fig. 8. Disparity of stereo images.



(a) Left image.

(b) Right image.

Fig. 9. Results of waterdrop removal.



(a) Distant scene.

(b) Close scene.

(c) Result of (a).

(d) Result of (b).

Fig. 10. Magnified improved left images.

Figure 12 shows the result when there are mud blobs on the protecting glass. In this experiment, the distance between two cameras must be very small by calculating the constraints of the camera configuration. It is impossible to realize a small baseline length in used stereo camera configuration because of the camera size. Therefore, we used a half mirror to solve this problem (Fig. 13). Figures 12(a)(b) are original left and right images, and Fig. 12(c) is a left image after chromatic registration. Figure 12(e)(f) are the final results of left and right images, while Fig. 12(d) is the result of left image without chromatic registration. From this result, the importance of the chromatic registration between left and right images when we use a half mirror that changes the color of the image can be verified. Of course, the accuracy of 3D measurement is not high in the case of small baseline length. However, the noise removal can be easily executed. This relationship is trade-off, and we must decide the baseline length according to the situations. Nevertheless, it is verified that our method can remove not only waterdrops but also mud blobs.

From these results, we can confirm the validity of the proposed noise removal method for a distant scene and a close range scene.

## VII. CONCLUSION

In this paper, we proposed a method for removing noises that disturb a view in stereo images. Our method is effective for removal of stationary noises in images that are difficult to remove by background subtraction or interframe subtraction in principle. Experimental results show the validity of noise removal for a close-range view stereo image pair that has disparities.

As a future work, we should improve the precision of disparity estimation. We have to also reduce a computation time and construct real-time processing method. It can be realized by using an image processing hardware when corresponding points between two images are detected. In addition, the chromatic registration method between left and right images must be sophisticated (*e.g.* [17]) for generating natural images. The combination of our method and physical based vision method (*e.g.* [18], [19]) is also our challenging future work.
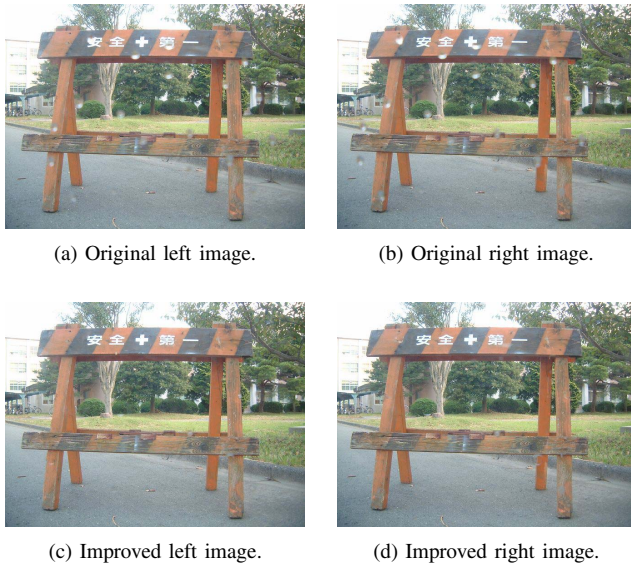
(a) Original left image.

(b) Original right image.



(c) Improved left image.

(d) Improved right image.

Fig. 11.   Result in the case of close scene.



(a) Original left image.

(b) Original right image.



(c) Color registration result of (a).

(d) Result without registration.



(e) Improved left image.

(f) Improved right image.

Fig. 12.   Result of mud blob removal.



(a) Parallel stereo camera.

(b) Stereo camera with half mirror.

Fig. 13.   Configuration of small baseline stereo camera.

## REFERENCES

[1] Anil C. Kokaram, Robin D. Morris, William J. Fitzgerald and Peter J. W. Rayner: "Detection of Missing Data in Image Sequences," *IEEE Transactions on Image Processing*, Vol.4, No.11, pp.1496–1508, 1995.

[2] Atsushi Nagai, Yoshinori Kuno and Yoshiaki Shirai: "Surveillance System Based on Spatio-Temporal Information," *Proceedings of the 1996 IEEE International Conference on Image Processing* (*ICIP1996*), Vol.2, pp.593–596, 1996.

[3] Hiroyuki Hase, Kazunaga Miyake and Masaaki Yoneda: "Real-time Snowfall Noise Elimination," *Proceedings of the 1999 IEEE International Conference on Image Processing* (*ICIP1999*), Vol.2, pp.406–409, 1999.

[4] Takashi Matsuyama, Takashi Ohya and Hitoshi Habe: "Background Subtraction for Non-Stationary Scenes," *Proceedings of the 4th Asian Conference on Computer Vision* (*ACCV2002*), pp.662–667, 2002.

[5] Kshitiz Garg, Shree K. Nayar: "Detection and Removal of Rain from Videos," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (*CVPR2004*), Vol.1, pp.528–535, 2004.

[6] Anil C. Kokaram, Robin D. Morris, William J. Fitzgerald and Peter J. W. Rayner: "Interpolation of Missing Data in Image Sequences," *IEEE Transactions on Image Processing*, Vol.4, No.11, pp.1509–1519, 1995.

[7] Simon Masnou and Jean-Michel Morel: "Level Lines Based Disocclusion," *Proceedings of the 5th IEEE International Conference on Image Processing* (*ICIP1998*), pp.259–263, 1998.

[8] L. Joyeux, O. Buisson, B. Besserer and S. Boukir: "Detection and Removal of Line Scratches in Motion Picture Films," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition* (*CVPR1999*), pp.548–553, 1999.

[9] Marcelo Bertalmio, Guillermo Sapiro, Vicent Caselles and Coloma Ballester: "Image Inpainting," *Computer Graphics* (*SIGGRAPH 2000*), pp.417–424, 2000.

[10] Sung Ha Kang, Tony F. Chan and Stefano Soatto: "Inpainting from Multiple Views," *Proceedings of the 1st International Symposium on 3D Data Processing Visualization and Transmission*, pp.622–625, 2002.

[11] Tony F. Chan and Jianhong Shen: "Variational Image Inpainting," *IMA Preprint*, 1868, pp.1–28, 2002.

[12] Atsushi Yamashita, Toru Kaneko and Kenjiro T. Miura: "A Virtual Wiper –Restoration of Deteriorated Images by Using a Pan-Tilt Camera–," *Proceedings of the 2004 IEEE International Conference on Robotics and Automation* (*ICRA2004*), pp.4724–4729, 2004.

[13] Atsushi Yamashita, Tomoaki Harada, Toru Kaneko and Kenjiro T. Miura: "Removal of Adherent Noises from Images of Dynamic Scenes by Using a Pan-Tilt Camera," *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS2004*), pp.437–442, 2004.

[14] Atsushi Yamashita, Tomoaki Harada, Toru Kaneko and Kenjiro T. Miura: "Virtual Wiper –Removal of Adherent Noises from Images of Dynamic Scenes by Using a Pan-Tilt Camera–," *Advanced Robotics*, Vol.19, No.3, pp.295–310, 2005.

[15] Atsushi Yamashita, Masayuki Kuramoto, Toru Kaneko and Kenjiro T. Miura: "A Virtual Wiper –Restoration of Deteriorated Images by Using Multiple Cameras–," *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS2003*), pp.3126–3131, 2003.

[16] S. G. Lipson, H. Lipson and D. S. Tannhauser: *Optical Physics Third Edition*, Cambridge University Press, 1995.

[17] Lisa Gottesfeld Brown: "A Survey of Image Registration Techniques," *ACM Computing Surveys*, Vol.24, No.4, pp.325–376, 1992.

[18] Srinivasa G. Narasimhan and Shree K. Nayar: "Vision and the Atmosphere," *International Journal of Computer Vision*, Vol.48, No.3, pp.233–254, 2002.

[19] Srinivasa G. Narasimhan and Shree K. Nayar: "Interactive (De)Weathering of an Image using Physical Models," *Proceedings of IEEE Workshop on Color and Photom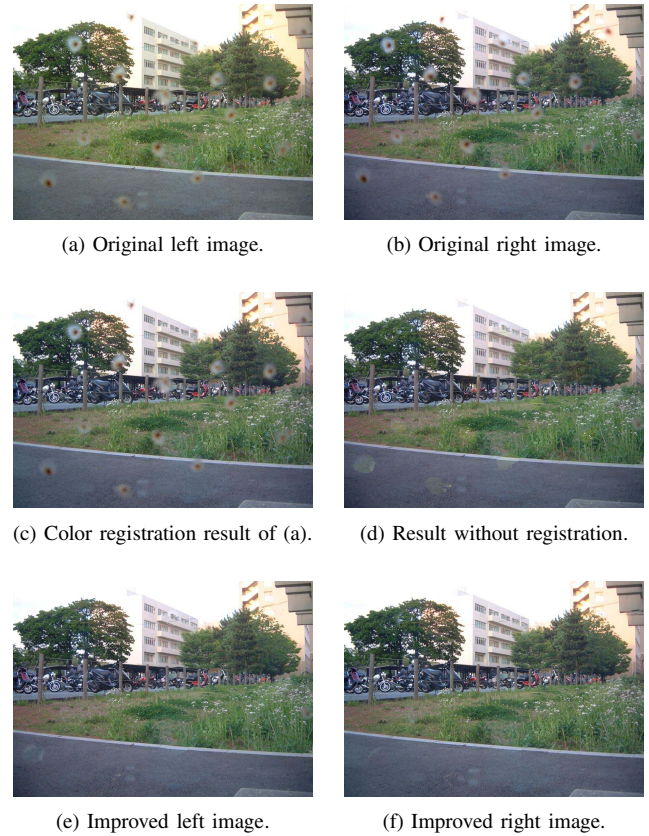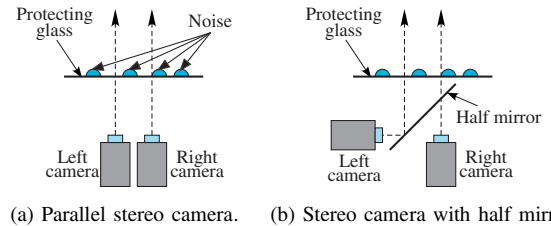etric Methods in Computer Vision* (*CPMCV*), 2003.