A Virtual Wiper
–Restoration of Deteriorated Images by Using a Pan-Tilt Camera–

Atsushi Yamashita*, Toru Kaneko* and Kenjiro T. Miura*
* *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 waterdrops from images by processing images taken with a pan-tilt camera system. In rainy days, it is often the case that images taken by the camera are hard to see because of adherent waterdrops on the surface of the protecting glass of the camera. In our method, an image of a distant prospect is taken at first and another image is taken after changing the direction of eyeshot. The new image is transformed with the projective transformation and compared with the first one to detect the region where waterdrops may exist. We can distinguish which image portion belongs to waterdrops by considering the distance between two waterdrop candidate regions. Finally, the region where waterdrops exist can be eliminated to merge two images. Experimental results show the effectiveness of the proposed method.**

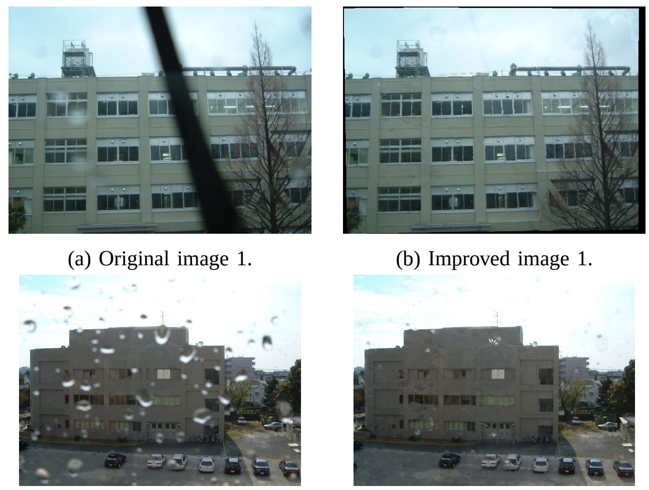
Index Terms*— *image restoration, noise removal, adherent waterdrop, pan-tilt camera, image composite

I. Introduction

Recently, it becomes very important to detect trespassers automatically by surveillance cameras in outdoor environments. However, in outdoor environments, it is often the case that scenes taken by the cameras are hard 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 block the visual field in rainy days. It would be desirable to remove adherent noises from images of such scenes for the surveillance and the environment recognition.

Therefore, we propose a new method for the restoration of deteriorated images. Specifically, we aim at removal of adherent waterdrops in images. The detection of noise areas in images and the interpolation of these 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], [2], [3], [4]. These techniques remove moving objects or noises by taking the difference between the initial background scene and a current scene, or taking the difference between temporarily adjacent two frames. These methods are robust against the change of background[2], the change of the weather[3], or the change of the lighting condition[4]. However, it is difficult to apply these techniques to the above problem, because adherent noises such as waterdrops may be stationary noises in the images.



(a) Original image 1. (b) Improved image 1. (c) Original image 2. (d) Improved image 2.

Fig. 1. Restoration of deteriorated images by using multiple cameras[10].

On the other hands, the image interpolation or restoration techniques for damaged and occluded images are also proposed[5], [6], [7], [8]. However, applying these methods require indicating the region of noises interactively (not automatically). It is also very difficult to treat large noises and to duplicate the complex textures with these methods.

Therefore, we propose the method for the removal of view-disturbing noises such as waterdrops from images taken with multiple cameras to treat this problem[9], [10] (Fig. 1). However, the situation where this method can be applied is limited because multiple cameras cannot necessarily be prepared.

In this paper, we propose a new method that can remove noises from images by using a pan-tilt camera. There are not few cases where a pan-tilt camera system is used for the surveillance. Therefore, we aim at solving a real-world problem and propose an effective solution to this problem.

Multiple images can be obtained when the camera rotates. The areas where noises exist are detected to compare two

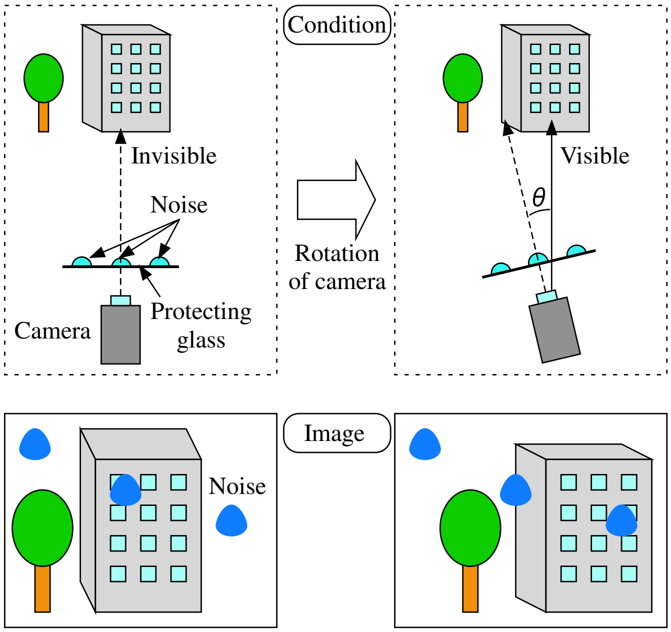


Fig. 2. Image acquisition.

images (the image before and after the camera rotation) and a clear image can be generated by combining clear regions of two images. This paper focuses on algorithms for distant view images.

The composition of this paper is detailed below. In Section II, the method of removing noises in images is explained. In Section III, experimental results are shown and Section IV discusses the effectiveness of our method. Finally, Section V describes conclusions and future works.

II. Image Restoration

The difference between images of same scene is very small where noises do not exist, and it is large where noises exist. On the other hand, the positions of waterdrops in images do not change when the direction of the camera changes (Fig. 2). This is because waterdrops are attached to the surface of the protecting glass of the camera and move together with the camera.

Therefore, we can obtain two images in which only the positions of waterdrops are different from each other when the image after the camera rotation is transformed to the image whose direction of eyeshot is same with that before the camera rotation. By taking into consideration the relationship of the waterdrops' positions in two images, the positions of waterdrops in each image can be estimated. Finally, the parts of images where no noises exist are merged to construct a clear image.

In this paper, the removal of waterdrops is realized by the following procedures.

1) Image Acquisition
2) Positional Registration
3) Chromatic Registration

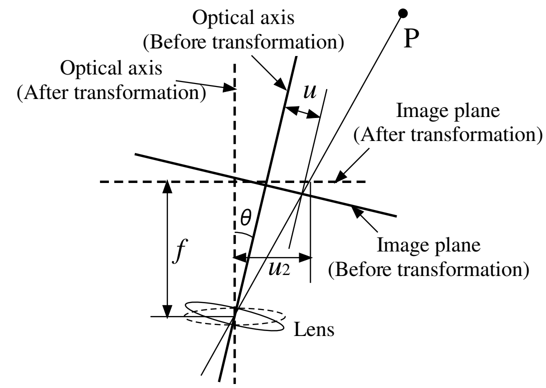


Fig. 3. Projective transformation.

4) Noise Region Extraction
5) Noise Judgment
6) Noise Removal

A. Image Acquisition

At first, one image is acquired where the camera is fixed. Next, another image is taken after the camera rotates θ degree about the axis which is perpendicular to the ground and passes along the center of the lens. The rotation angle θ makes a positive direction a counterclockwise rotation (the direction of Fig. 2).

B. Positional Registration

As to the positional registration, the distortion of two images is rectified, the image after the camera rotation is transformed by using the projective transformation, and two images of same scene in which only positions of waterdrops differs from each other are obtained.

At first, the distortion of two images is rectified. Let (u, v) be the coordinate value without distortion, (u', v') be the coordinate value with distortion (observed coordinate value), and κ_1 be the parameter of the radial distortion[11]. The distortion of the image is corrected by (1), (2).

$$u' = u + \kappa_1 u(u^2 + v^2), \quad (1)$$
$$v' = v + \kappa_1 v(u^2 + v^2). \quad (2)$$

Next, the image after the camera rotation is transformed by using the projective transformation. The coordinate value after transformation (u_2, v_2) is expressed as follows (Fig. 3):

$$u_2 = f \frac{f \tan\theta + u}{f - u \tan\theta}, \quad (3)$$
$$v_2 = f \frac{\sqrt{1 + \tan^2\theta}}{f - u \tan\theta} v, \quad (4)$$

where (u, v) is the coordinate value before transformation, θ is the rotation angle of the camera, and f is the image distance[1].

[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[12].

The image after the camera rotation is transformed to the image whose direction of eyeshot is same with that before the camera rotation.

C. Chromatic Registration

It is often the case that the chromatic tone of the image after the camera rotation changes from that before the camera rotation under the influence of the change of the lighting condition. Therefore, chromatic registration about the common field of view of two images must be executed after the positional one.

We express the relationship of the characteristics of the color reproduction between two images as the linear function and match the RGB values of each image with these functions.

Here, let $R_0(u,v)$, $G_0(u,v)$, $B_0(u,v)$ be the RGB value of the initial image before the camera rotation, respectively. In the same way, let $R_\theta(u,v)$, $G_\theta(u,v)$, $B_\theta(u,v)$ be the RGB value of the image after the camera rotation, respectively.

The difference between two images changes greatly according to the levels of the pixel value in images. Therefore, the whole range of RGB value (0–255) is divided into several classes, and the average of RGB value of two images in the same class is fit with each other.

The RGB difference between two images in each class ($\overline{r_c}$, $\overline{g_c}$, $\overline{b_c}$) is calculated in (5)–(7), respectively.

$$\overline{r_c} = \frac{1}{m_{r,c}} \sum_{u=u_s}^{u_e} \sum_{v=v_s}^{v_e} \{R_{\theta,c}(u,v) - R_{0,c}(u,v)\}, \quad (5)$$

$$\overline{g_c} = \frac{1}{m_{g,c}} \sum_{u=u_s}^{u_e} \sum_{v=v_s}^{v_e} \{G_{\theta,c}(u,v) - G_{0,c}(u,v)\}, \quad (6)$$

$$\overline{b_c} = \frac{1}{m_{b,c}} \sum_{u=u_s}^{u_e} \sum_{v=v_s}^{v_e} \{B_{\theta,c}(u,v) - B_{0,c}(u,v)\}, \quad (7)$$

where u_s, u_e, v_s, v_e is (u,v) coordinate value about the common field of view of two images, c is class number, and $m_{r,c}$, $m_{g,c}$, $m_{b,c}$ is total sum of pixel of RGB component in each class, respectively.

In the next step, the new RGB value after the chromatic registration $R'_{\theta,c}(u,v)$, $G'_{\theta,c}(u,v)$, $B'_{\theta,c}(u,v)$ is calculated in (8)–(10).

$$R'_{\theta,c}(u,v) = R_{\theta,c}(u,v) + \overline{r_c}, \quad (8)$$

$$G'_{\theta,c}(u,v) = G_{\theta,c}(u,v) + \overline{g_c}, \quad (9)$$

$$B'_{\theta,c}(u,v) = B_{\theta,c}(u,v) + \overline{b_c}. \quad (10)$$

D. Noise Region Extraction

The positions where noises exist are estimated by comparing two images. Here, it should be noted that we use monochromatic gray-scale images converted from the color images obtained above. The difference between two gray-scale images about the common field of view of two images is calculated, and the thresholding process gives a difference image where noise regions and the rest are binarized (Fig. 4). Here, let $I_0(u,v)$ and $g_1(u,v)$ be a color and a gray-scale image before

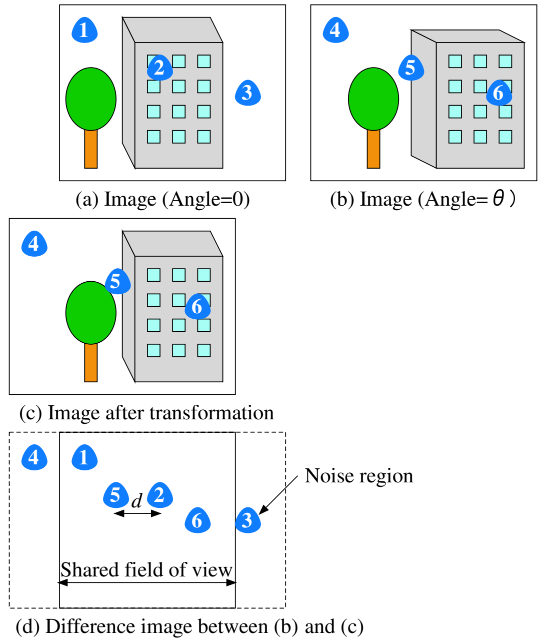


Fig. 4. Projective transformation and difference between two images.

the camera rotation, respectively, and $I'_\theta(u,v)$ and $g_2(u,v)$ be a transformed color and a gray-scale image after the camera rotation. We define regions where the differences between two gray-scale images are larger than a certain threshold T as the noise regions of two images. The difference image $G(u,v)$ is obtained by

$$G(u,v) = \begin{cases} 0, & |g_1(u,v) - g_2(u,v)| \leq T \\ 1, & |g_1(u,v) - g_2(u,v)| > T \end{cases}. \quad (11)$$

The region of $G(u,v) = 1$ is defined as noise regions. However, the regions where no waterdrops exist are extracted in this process because of image noises. Therefore, the morphological operations (opening operation: erosion and dilation) are executed for eliminating small noises. After that, the size of each noise region N_i is calculated.

E. Noise Judgment

It is judged to which image each noise region is attached between the image before the camera rotation and the transformed image after the camera rotation. The waterdrop attached to the transformed image looks moving leftward to the same waterdrop attached to the image before the camera rotation when the direction of the camera rotation is counter-clockwise ($\theta \geq 0$). The distance between two waterdrops d is defines as follows:

$$d(u_i) = \frac{(f^2 + u_i^2)\tan\theta}{f - u_i \tan\theta}, \quad (12)$$

where u_i is u-coordinate value of the representative point of attached waterdrop.

It can be judged to which image each noise region is attached by distinguishing whether there are the same size noise region which is distant from another noise region and

TABLE I

RULE OF NOISE JUDGMENT.

case	1	2	3	4	5	6	7	8	9
left	O	O	O	X	X	X	–	–	–
right	O	X	–	O	X	–	O	X	–
noise	*1	0	0	θ	*2	θ	θ	0	*3

whose distance from there is d. The conditions of regarding the size of two noise region as same is expressed as follows:

$$\frac{\min(N_1, N_2)}{\max(N_1, N_2)} > U, \tag{13}$$

where N_1 is the size of one noise region and N_2 is the size of the noise region whose distance from the above region is d.

The judgment to which image each noise region is attached is executed from the rule about the distance and the sizes of two noise regions shown in Table I.

When a certain noise region is observed, three cases can be considered: there is a noise region whose distance from the observed noise region is d (O in Table I), there is not a noise region (X in Table I), and the place whose distance from the observed noise region is d is outside of the image (– in Table I). Since three cases can be considered about rightward and leftward of the noise region, there are nine cases about the relationship between the noise regions whose distance is d.

About the noise judgment, “0” means that the noise region is attached to the image before the camera rotation, and “θ” means that that is attached to the image after the projective transformation in Table I.

For example, in Fig. 4(d), there is a waterdrop in the left hand of waterdrop 2, and there is no waterdrop in the right hand of it. Therefore, waterdrop 2 corresponds to the case 2 in Table I, and waterdrop 2 is judged to be attached to the image before the camera rotation. In the same way, the image in which there is a noise region can be judged in cases 3, 4, 6, 7, and 8.

Here, the case in which the judgment cannot be executed is considered. In case 1 (*1 in Table I), there are same size noise regions in the left and right hands of the observed noise region. In case 5 (*2 in Table I), there are not same size noise regions in the left and right hands of the observed noise region. This is because two waterdrops overlaps when the difference image is calculated, or there exists moving objects in the scene. When two waterdrops overlap, it is impossible to remove these waterdrops.

In case 9 (*3 Table I), the places in the left and right hands of the noise region are outside of image. This is the case that the rotation angle of the camera is very large.

In this paper, we decide to judge that there is waterdrops in the base image of final result when we cannot judge in which image waterdrop exists (cases 1, 5, and 9).

In addition, when the direction of camera rotation is clockwise ($\theta < 0$), the rules of judgment become contrary to Table I.

F. Noise Removal

Waterdrop removal is performed by using another image data for the noise regions. The image composition of two images is expressed as follows:

$$\begin{aligned} F(u,v) &= \alpha(u,v) I_0(u,v) \\ &\quad + (1-\alpha(u,v)) I'_\theta(u,v), \end{aligned} \tag{14}$$

where $0 \leq \alpha(u,v) \leq 1$.

It occurs that the removal of waterdrop contour cannot be performed because the difference between the area around the contour and the scenery is small and then the noise region becomes small. Therefore, the dilation operation is executed about each noise region.

In this paper, the final image $F(u,v)$ is made based on the image before the camera rotation $I_0(u,v)$. The transformed image $I'_\theta(u,v)$ is used in the region where the waterdrops exist in $I_0(u,v)$.

The image $G'(u,v)$ is finally constructed by removing the noise regions attached to the transformed image $I'_\theta(u,v)$ from the difference image $G(u,v)$. We can set $\alpha(u,v) = 1$ when $G'(u,v) = 0$, and $\alpha(u,v) = 0$ when $G'(u,v) = 1$.

Additionally, if the chromatic tones of two images differ from each other, a delicate chromatic difference may remain after the chromatic registration. Therefore, we change $\alpha(u,v)$ value around the noise region linearly from 0 to 1, because of natural composition of images.

G. Rotational Angle of Camera

It is desirable that the rotational angle of the camera is small because of large common field of view of two images. However, it is necessary for the same waterdrop not to overlap in two images before and after the camera rotation.

When the size of the biggest waterdrop in lateral direction is equal to S, $d > S$ must be satisfied. Therefore, the constraints about the rotational angle of the camera are expressed as (15).

$$\theta > \tan^{-1}\left(\frac{S}{f}\right). \tag{15}$$

III. EXPERIMENTS

At first, the camera calibration was done. When the resolution of the camera is 640×480pixel, it was estimated that $f = 764.6$pixel and $\kappa_1 = -2.67 \times 10^{-7}$.

If $S = 100$pixel, the rotational angle of the camera θ must be larger than 7.5deg from (15). Therefore, we set θ as 10.0deg because a margin is needed for safety’s sake.

About the chromatic registration, we divided the whole range of RGB value (256 level) into 8 classes at regular intervals. All parameters were set by the trial-and-error method. In our experiment, we set that $T = 20$, $U = 0.65$.

Figure 5–9 show the results of our method. Figure 5 shows the image before the camera rotation ($\theta = 0$deg), Fig. 6 shows the image after the camera rotation ($\theta = 10$deg), Fig. 7 shows the transformed image after the camera rotation, Fig. 8 shows the positions of waterdrops attached to the image before the camera rotation, and Fig. 9 shows the result image of waterdrop removal. This result indicates that our method can remove waterdrops with propriety.

About the quantitative evaluation, 92.6% of the noises in the original image can be removed in the improved image (Fig. 9). Remaining 7.4% noises were on the part of image where the difference between two images is very small. Actually, the waterdrops that were not able to be removed can hardly be recognized, and a clear image can be gained.

Fig. 5. Image before the change of camera's direction (Initial image).

Fig. 6. Image after the change of camera's direction (Rotation image).

Figure 10 shows a part of another image. Figure 10(a) shows the original image in which waterdrop attached on the edge of the background object, Fig. 10(b) is the result by our proposed method. A result with "image inpainting" algorithm [6] is shown in Fig. 10(d). In the case of image inpainting, human operator indicates the position of adherent waterdrop (Fig. 10(c)). Figure 10(b) is correctly restored, although the edge of the background object is not correctly restored in Fig. 10(d)[2].

[2]Note that all parameters of "image inpainting" algorithm [6] were not perfectly set correctly in our experiments.

Fig. 7. Image after the projective transformation (Transformed image).

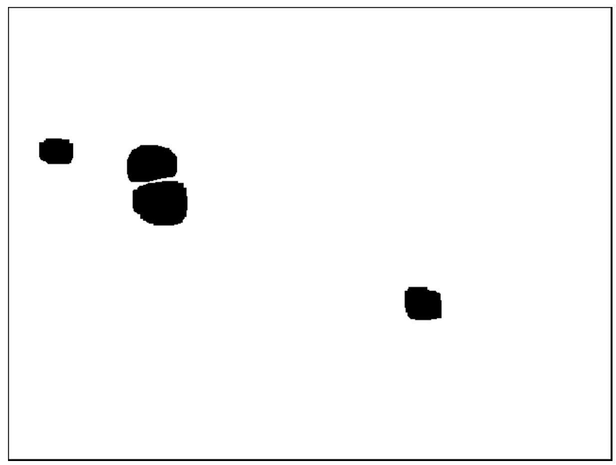


Fig. 8. Waterdrops in the initial image.

From these results, it is verified that clear images can be obtained in the conditions that waterdrops do not overlap between two images.

IV. Discussions

About the lighting conditions, we verified our method in the case of bright backgrounds, in the case of dark backgrounds, in the case that the average difference between two images is larger than 10 (the whole range: 256 level), and it has been shown that our method could remove noises properly. Therefore, our method is independent from the lighting condition when it is daytime.

However, it is expected that our method cannot remove noises when the difference between two images is very large, or when a certain parts of the image drastically changes and separates from the average difference of the whole image greatly.

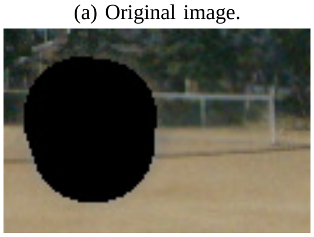
About the size of noises that can be removed, our method can eliminate noises whose size is from 100pixel to 5500pixel.

As to the rotation of the camera, the position of principle center of the camera must correspond to that of the rotational

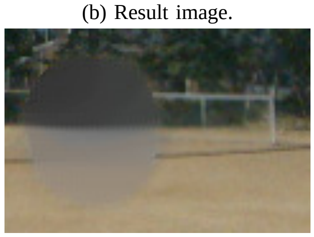
Fig. 9. Image after removing waterdrops.

(a) Original image.

(b) Result image.



(c) Waterdrop area.



(d) Result image by [6].

Fig. 10. Example of waterdrop removal.

center of the camera, although it is very difficult to fit them. It was expected that the disagreement occurred in our experiment. However, the influence of the disagreement between the principle point and the rotational center is very small as compared with the influence of other factors such as the image distortion, and it can be disregard. Actually, there was no positional difference between two images after the positional registration.

Furthermore, our proposed method in this paper is suitable for high-speed processing as compared with the method that uses multiple cameras[10]. This is because this method can judge noises by simple right-and-left comparison, although the method by using multiple cameras needs to calculate several features in images.

V. Conclusions

In this paper, we propose a new method that can remove view-disturbing waterdrops from images by processing images taken with a pan-tilt camera system that can change the direction of eyeshot. In our method, an image of a distant prospect is taken at first and another image is taken after changing the direction of eyeshot. The new image is transformed with the projective transformation and compared with the first one to detect the region where waterdrops may exist. The region where waterdrops exist can be eliminated to merge two images. Experimental results show the effectiveness of the proposed method.

The proposed method is simple and effective, and therefore it has the potential to be used in several places.

As the future works, an automatic determination of threshold values and an application to movies are needed. The chromatic registration method in this paper must be sophisticated, e.g.[13].

Acknowledgment

This research was partially supported by Special Project for Earthquake Disaster Mitigation in Urban Areas (in cooperation with International Rescue System Institute (IRS) and National Research Institute for Earth Science and Disaster Prevention (NIED)), the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Young Scientists (B), 15700153, 2003, and the Sawagawa Scientific Research Grant from the Japan Science Society.

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] 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.

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

[7] 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.

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

[9] Masayuki Kuramoto, Atsushi Yamashita, Toru Kaneko and Kenjiro T. Miura: "Removal of Waterdrops in Images by Using Multiple Cameras," *Proceedings of IAPR Workshop on Machine Vision Applications* (*MVA2002*), pp.80–83, 2002.

[10] 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.

[11] Juyang Weng, Paul Cohen and Marc Herniou: "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.10, pp.965–980, 1992.

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

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