A Virtual Wiper
– Restoration of Deteriorated Images by Using Multiple Cameras –

Atushi Yamashita†, Masayuki Kuramoto†,‡, Toru Kaneko† and Kenjiro T. Miura†
† Department of Mechanical Engineering, Shizuoka University
3-5-1 Hamamatsu-shi, Shizuoka 432-8561, Japan
‡ Fuji Photo Film Co., Ltd.
E-mail address: yamashita@ieee.org

***Abstract*— In this paper, we propose a new method for the restoration of deteriorated images by using multiple cameras. In outdoor environment, 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 cameras. Our proposed method analyses multiple camera images describing the same scene, and synthesizes an image in which adherent noises are eliminated.**

key words: adherent noise, noise elimination, image restoration, image compositing, multiple cameras

I. INTRODUCTION

Recently, it becomes very important to detect trespassers automatically by surveillance cameras in outdoor environments. The task that mobile robots collect the information about the environment by using cameras also will become very significant and be in high demand for security or disaster response in the near future. 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. 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.

Professional photographers use lens hoods or put special water-repellent oil on lens to avoid this problem. Even in these cases, waterdrops are still attached on the lens. Cars are equipped with windscreen wipers to wipe rain from their windscreens. However, there is a problem that a part of the scenery is not in sight when a wiper crosses.

Therefore, we propose a new method for the restoration of deteriorated 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]–[4]. These techniques remove the 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. 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.

On the other hand, the image interpolation or restoration techniques for damaged and occluded images are also proposed [5]–[8]. However, applying these methods require to indicate 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.

In this paper, we propose a new method for the removal of view-disturbing noises from images taken with multiple cameras (Fig. 1). Our method extracts the positions of adherent noises by comparing multiple images, and merges the parts of images where noises do not exist. This paper focuses on algorithms for the restoration of deteriorated images from multiple images of a distant scene, in which little stereoscopic disparities exist. We construct the methods for a stereo camera system, and for a three-camera system when a large number of noises exist.

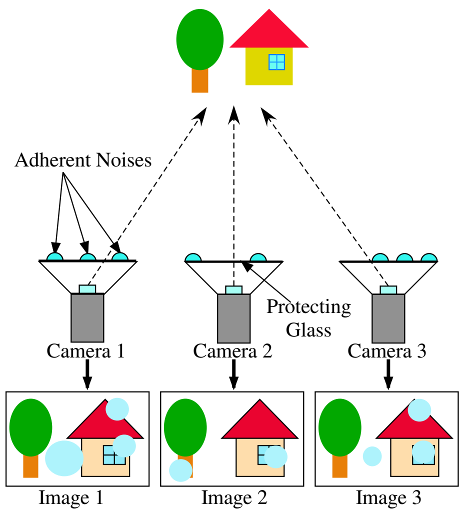


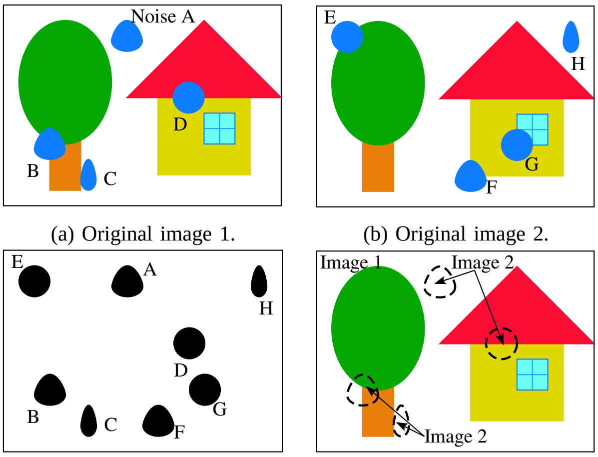
Fig. 1. Image acquisitions.

II. RESTORATION OF DETERIORATED IMAGES

In this paper, it is assumed that the distances between cameras are quite smaller than the distance of scenery. When images of a scene are taken with two or three cameras, the difference between images is very small where noises do not exist, and it is large where noises exist (Fig. 2(a)(b)). The region of the noises can be extracted by using the difference between two images (Fig. 2(c)). This region itself, however, does not have information in which image noises exist. Therefore, we estimate in which image each noise attaches by using the features of noises in images and the set operations. Finally, the parts of images where no noises exist are merged to construct a clear image (Fig. 2(d)).

The procedure for the restoration of deteriorated images consists of four steps:

1) image registration
2) extraction of noise regions
3) judgment of noise regions
4) noise removal



(a) Original image 1. (b) Original image 2.

(c) Noise regions. (d) Image restoration.

Fig. 2. Overview of our proposed method.

III. IMAGE REGISTRATION

At the first step, two or three images of a scene are acquired simultaneously by using multiple cameras. Since it is very important for the extraction of noises to take the exact difference between image pairs, positional and chromatic registrations are needed.

Generally, the images from the cameras are distorted under the influence of the lens aberration. Therefore, the distortion of the image is corrected at first [9].

Secondly, the positional registration between images is executed. Ideally, background images are same when multiple cameras are set up that each camera's optical axis and its scanline are parallel to each other, respectively. However, the background images are different from each other and do not describe the same scene according to the setting error of cameras and the delicate difference between them. Therefore, the positional registration is executed with the projective transformation (1)–(2).

$$u_{new} = \frac{a_{11}u + a_{12}v + a_{13}}{a_{31}u + a_{32}v + 1}, \quad (1)$$

$$v_{new} = \frac{a_{21}u + a_{22}v + a_{23}}{a_{31}u + a_{32}v + 1}, \quad (2)$$

where (u, v) are original coordinate values, (u_{new}, v_{new}) are coordinate values after transformation, and a_{ij} are coefficients. The projective transformation can fit the positions of planes between images, and it can be used in the case of the images of a distant scene. We can obtain a_{ij} by finding at least four corresponding points between image pairs.

Finally, the chromatic registration is achieved. We express the relationship of the characteristics of the color reproduction between two cameras as the linear function and match the RGB values of each image with these functions.

IV. EXTRACTION OF NOISE REGIONS

A. Difference Image

At the second step, 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. We define regions where the differences between two gray-scale images are larger than a certain threshold as the noise regions of two images. The difference between two images is calculated, and the thresholding process gives a difference image where noise regions and the rest are binarized. The difference image $g_{ij}(u, v)$ is obtained by

$$g_{ij}(u,v) = \begin{cases} 0, & |f_i(u,v) - f_j(u,v)| \leq L(u,v) \\ 1, & |f_i(u,v) - f_j(u,v)| > L(u,v) \end{cases}, \quad (3)$$

where $i,\ j\ (= 1, 2, 3)$ are image numbers, $f_i(u, v)$ is the pixel value of the i-th gray-scale image at pixel (u, v), and $L(u, v)$ is a threshold value. The region of $g_{ij}(u, v) = 1$ is defined as noise regions. Hereafter, we call the difference image of the noise regions as a NR image (Fig. 2(c)).

The threshold value $L(u, v)$ should be determined for each noise region independently, because its optimum value differs from each other. If the threshold value is too large, the size of the noise region becomes smaller than the actual size, or the noise region vanishes. If the threshold is too small, the size of the noise region becomes larger than the actual size, and many noises appear as false noise regions. Therefore, it is necessary to determine the appropriate threshold value automatically for each noise region by using the features of noise regions in images.

B. Feature of Adherent Noises

There exist the following two features about the noise regions:

1) The pictures of the noise regions are blurred. This means that the change (variance) inside the noise region of the picture that contains the noise becomes small as compared with the pictures without the noise.
2) The contour (edge) of the noise region outstands. This means that the change (variance) at the contour pixel of the noise region of the picture that contains the noise becomes large as compared with the pictures without the noise.

The change (variance) inside the noise region $I_{i,l}$ is expressed as follows:

$$I_{i,l} = \frac{1}{h_l} \sum_{(u,v)\in R_l} \left\{ f_i(u,v) - \frac{1}{h_l} \sum_{(u,v)\in R_l} f_i(u,v) \right\}^2, \qquad (4)$$

where i is an image number, l is a labeled number of the noise regions, R_l is a set of the pixel inside the noise region l, h_l is a total number of pixel in R_l.

The change (variance) at a contour pixel is given by the sum of squared values of the difference between each pixel within the 3×3 pixel window around the contour pixel and the mean value in the gray-scale image $f_i(u,v)$. Then we calculate the average variance $C_{i,l}$ along the pixels belonging to one contour, by the following equation:

$$\begin{aligned} V_{i,k} &= \frac{1}{9} \sum_{u=\alpha_{k,l}-1}^{\alpha_{k,l}+1} \sum_{v=\beta_{k,l}-1}^{\beta_{k,l}+1} \Bigg\{ f_i(u,v) \\ &\quad - \frac{1}{9} \sum_{u=\alpha_{k,l}-1}^{\alpha_{k,l}+1} \sum_{v=\beta_{k,l}-1}^{\beta_{k,l}+1} f_i(u,v) \Bigg\}^2, \qquad (5) \\ C_{i,l} &= \frac{1}{n_l} \sum_{k=1}^{n_l} V_{i,k}, \qquad (6) \end{aligned}$$

where $V_{i,k}$ is the variance at k-th contour pixel of l-th noise region, $(\alpha_{k,l}, \beta_{k,l})$ is the pixel coordinate value belonging to the contour, and n_l is the number of pixels belonging to the contour.

C. Decision of Threshold Value

The thresholding method is based on the feature 2) that a noise has an edge contour against the background image and that a variance of pixel values along this contour $C_{i,l}$ is large compared to that along inner contours within a noise. The appropriate threshold value $L_l(u,v)$ of l-th noise region is decided by the exploratory search while changing $L_l(u,v)$ for every noise regions. $L_l(u,v)$ that satisfies (7) is regarded as the optimal threshold value of the l-th noise region.

$$C_{kl}(L_l(u,v)) \rightarrow max. \qquad (7)$$

After finding the optimal $L_l(u,v)$ and binarizing difference image, the morphological operations (the contraction and expansion operations) are executed for eliminating small noises of the NR images.

V. Judgment of Noise Regions

The NR images themselves have no information in which image noises exist. Therefore, noises are distinguished by combining the feature values of the noise regions when the number of images is two. The set operations can be used in addition when the number of images is three for the situation that there are a lot of adherent noises in images and that the noises exist in the same place of images.

A. Judgment from Two Images

We can distinguish the noise regions by using the two features mentioned in Section IV-B.

When the background texture of images is simple, the variance of the image that contains a noise at a contour pixel of the noise region (Fig. 3(a)) is larger than that of the image without the noise (Fig. 3(b)). Therefore, the feature 2) can be used in the area where the background texture is simple, and it can be judged that the image with the larger $C_{i,l}$ contains the noise region.

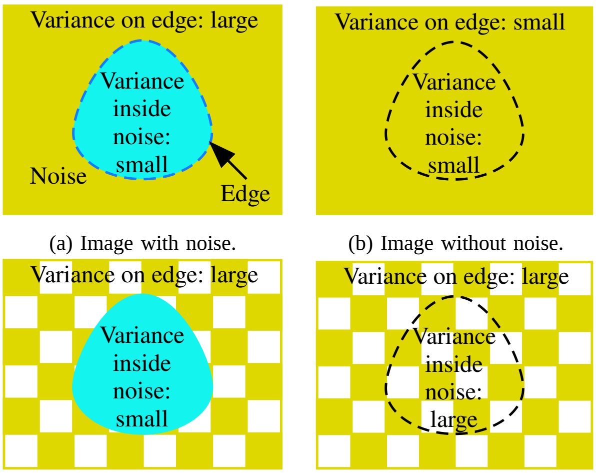
When the background texture of images is complicated, the variances of the two images at a contour pixel of the noise region do not differ from each other, because the variances become large on the complicated background whether the noise exists or not. In this case, the variance inside the noise region of the image that contains the noise (Fig. 3(c)) is smaller than that of the image without the noise (Fig. 3(d)). Therefore, the feature 1) can be used in the area where the background texture is complicated, and it can be judged that the image with the smaller $I_{i,l}$ contains the noise region.

The complexity of the background texture where noise regions exist can be checked by using the variance inside the noise region. The background texture is regarded simple when $I_{1,l} < P$ and $I_{2,l} < P$. That is complicated when $I_{1,l} > P$, $I_{2,l} > P$, and $|I_{1,l} - I_{2,l}| > Q$ (P, Q: threshold values given in advance).

When the background is neither simple nor complicated, the average pixel value inside the noise region $\overline{f_{i,l}}$ can be used. As to the noises such waterdrops, the average pixel value inside the noise is very high (white) because waterdrops on the protecting glass condense the light. As to the other noises, this value is very low (black) because these noises look like colored blobs. Therefore, it can be judged that the image with the larger $\overline{f_{i,l}}$ contains the noise region when $(\overline{f_{1,l}} + \overline{f_{2,l}})/2 > R$, and that with the smaller $\overline{f_{i,l}}$ contains the noise region when $(\overline{f_{1,l}} + \overline{f_{2,l}})/2 \leq R$ (R: a threshold value given in advance).

B. Judgment from Three Images

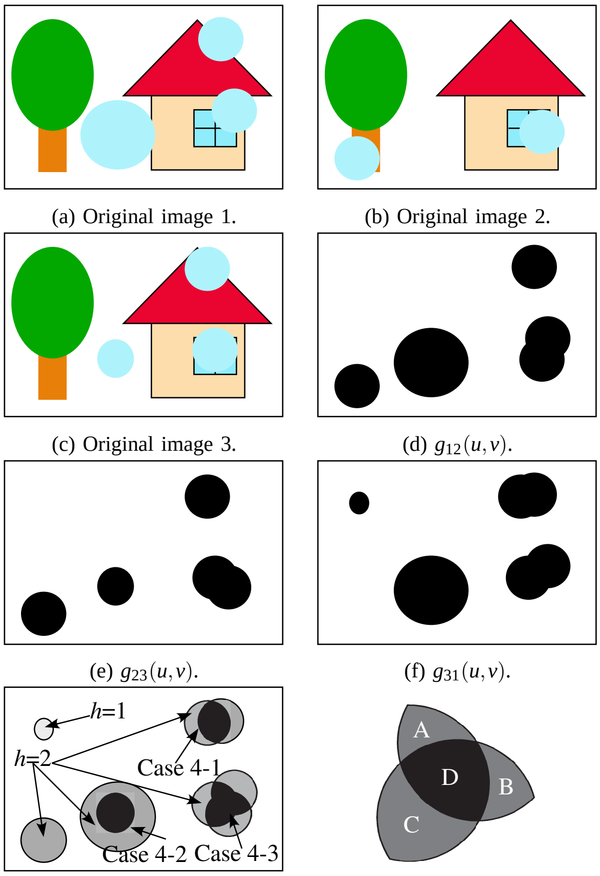
When there are three images, a simple majority decision method may be used. However, it cannot be used because a wrong judgment is executed when adherent noises exist at the same place on two images. Therefore, the noise



(a) Image with noise. (b) Image without noise.

(c) Image with noise. (d) Image without noise.

Fig. 3. The features of the noise region. (a)(b): Simple background case. (c)(d): Complicated background case.



(a) Original image 1. (b) Original image 2.

(c) Original image 3. (d) $g_{12}(u,v)$.

(e) $g_{23}(u,v)$. (f) $g_{31}(u,v)$.

(g) $h(u,v)$. (h) Case4-3.

Fig. 4. Judgment of noise regions.

regions are judged by combining the information of three NR images (8), and the set operations in addition to the feature values of three gray-scale images.

$$h(u,v) = \sum_{i \neq j} g_{ij}(u,v). \tag{8}$$

The value of $h(u,v)$ varies as 0, 1, 2, 3 (Fig. 4). When $h(u,v) = 0$, 1, and 2, the distinction is realized by a pixel-based processing.

Case 1: $h(u,v) = 0$

Noises do not exist in any of three images. Then all images can be judged that there are no noises.

Case 2: $h(u,v) = 1$

It is the case that the noise region exists between two images although the adherent noise does not exist in the original three images because of the setting of $L(u,v)$ or the individual difference of the cameras. When $g_{12}(u,v) = 1$, the pixel in the original image 1 or 2 belongs to a noise. Then the original image 3 can be used as a noise-free image. In the same way, the original image 1 and 2 can be used when $g_{23}(u,v) = 1$ and $g_{31}(u,v) = 1$, respectively.

Case 3: $h(u,v) = 2$

A noise exists only in one image of the three. When $g_{12}(u,v) = 0$, the pixels in the original image 1 and 2 do not belong to a noise. Similarly, either the original image 2 or 3 can be used when $g_{23}(u,v) = 0$, and either the original image 1 or 3 can be used when $g_{31}(u,v) = 0$.

Case 4: $h(u,v) = 3$

In this case, it is impossible to distinguish which pixel among three images belongs to a noise by the pixel-based judgment. The distinction is realized by the following region-based one. Here, we have three sub cases shown in Fig. 4(g).

Case 4-1: The region satisfying $h(u,v) = 3$ is surrounded by the two regions satisfying $h(u,v) = 2$ (Case 4-1 in Fig. 4(g)). The surrounding regions determine in which image a noise exists. Since the region satisfying $h(u,v) = 3$ belongs to a noise of the surrounding regions, we use the image in which the noise of the two surrounding regions does not exist.

Case 4-2: The region satisfying $h(u,v) = 3$ coincides with one of the noise region (Case 4-2 in Fig. 4(g)). The region satisfying $h(u,v) = 3$ in either of two images belongs to a noise. The image with the smaller $C_{i,l}$ is used as a noise-free image.

Case 4-3: The regions in three images corresponding to the region satisfying $h(u,v) = 3$ (Case 4-3 in Fig. 4(g)) contain pixels belonging and not belonging to a noise. In this case, the region-based judgment cannot be done because we cannot divide the region A, B, C, D in Fig. 4(h) by the set operations. Therefore, we use the pixel-based approach in this case. We find the latter pixels (A, B, C in Fig. 4(h)) by minority decision, because two images have a noise in the corresponding regions. The former pixels (D in Fig. 4(h)) belong to a noise in three images simultaneously and it is impossible to obtain a noise-free image.

VI. NOISE REMOVAL

The areas where no adherent noises exist are combined to make a clear image. For the natural image compositing, the morphological operation that expands the noise regions is used because there are possibilities that smaller size of the noise region than the real size is extracted.

VII. EXPERIMENTS

In the experiments, digital cameras are used for taking images. The resolutions of images are 640×480 pixel. All improved images are constructed on the basis of the original image 2.

Figure 5 shows the result by using two images that contain waterdrops as the adherent noises. The original images from two cameras are shown in Fig. 5(a)(b) and the image 1 after the registration process is shown in Fig. 5(c). The method gives a clearer image than original ones (Fig. 5(d)). The average occupation rate of noises in two original images to all area is 3.8% and that in the result image is 1.1%.

Figure 6 shows the result by using three images that contain waterdrops and black adherent noises. This result shows that our method can treat with the adherent noises independently of their colors. The average occupation rate in original images is 6.4% and that in the result image is 1.0%.

Figure 7 shows the result by using three images that contain waterdrops and a long shape noise like a windscreen wiper. This shows that a virtual wiper can remove an actual one and work independently of the noise shapes. The average occupation rate in original images is 6.9% and that in the result image is 0.2%.

Figure 8 shows the result by using three images that contain large number of noises such as waterdrops. The average occupation rate in original images is 12.3%. A result with two images (from Fig. 8(a)(b)) is shown in Fig. 8(d). This result is not clear because adherent noises exist at the same place on two images. A result with "image inpainting" algorithm [8] is shown in Fig. 8(f). In this case, human operator indicates the position of adherent noises (Fig. 8(e)). The result image is mostly restored, however, the position where large noises exist and the edge of the background cannot be restored finely[1]. A result with three images is shown in Fig. 8(h) and a result with a simple majority decision method is shown in Fig. 8(g) for a comparison. The proposed method gives a better result than a simple majority decision. The occupation rate is 0.8% with the proposed method, and 2.7% with the simple majority decision method.

[1]Note that all parameters of "image inpainting" algorithm [8] were not perfectly set correctly in our experiments.

(a) Original image 1.

(b) Original image 2.

(c) Registration result.

(d) Improved image.

Fig. 5. Experimental results I.

(a) Original image 1.

(b) Original image 2.

(c) Original image 3.

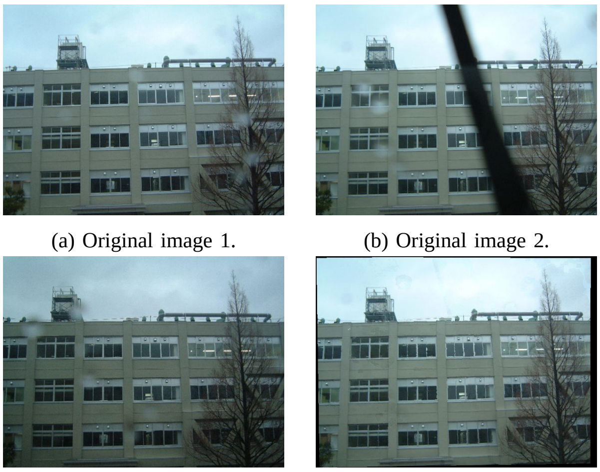
(d) Improved image.

Fig. 6. Experimental results II.

From these results, it is verified that our method can remove adherent noises without reference to the colors, the sizes, or the occupation rates of them.

VIII. CONCLUSIONS

In this paper, we proposed an effective method for the removal of view-disturbing noises adherent to the lens protecting glasses of cameras. The method is applied to two or three images of a distant scene, in which little stereoscopic disparities exist. The judgment of noises is realized by using the features of the noise regions and the set operations. The experimental results have shown the effectiveness of the proposed method. In principle, our method can treat the situation that noises / backgrounds are moving / stationary, and apply to moving / still images. While several ideas are employed in traditional image



(a) Original image 1. (b) Original image 2.

(c) Original image 3. (d) Improved image.

Fig. 7. Experimental results III.

restoration techniques, the size of defects that our method can deal with exceeds traditional techniques for digital scratch removal from films. Therefore, this method can be used at several situations in outdoor environments, e.g. surveillance, information collection by mobile robots, and so on.

As the future works, images in which large stereoscopic disparities exist is treated by using the correlation-based method when extracting the noise regions, while this method is based on difference image because of little stereoscopic disparities.

IX. ACKNOWLEDGMENTS

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Young Scientists (B), 15700153, 2003.

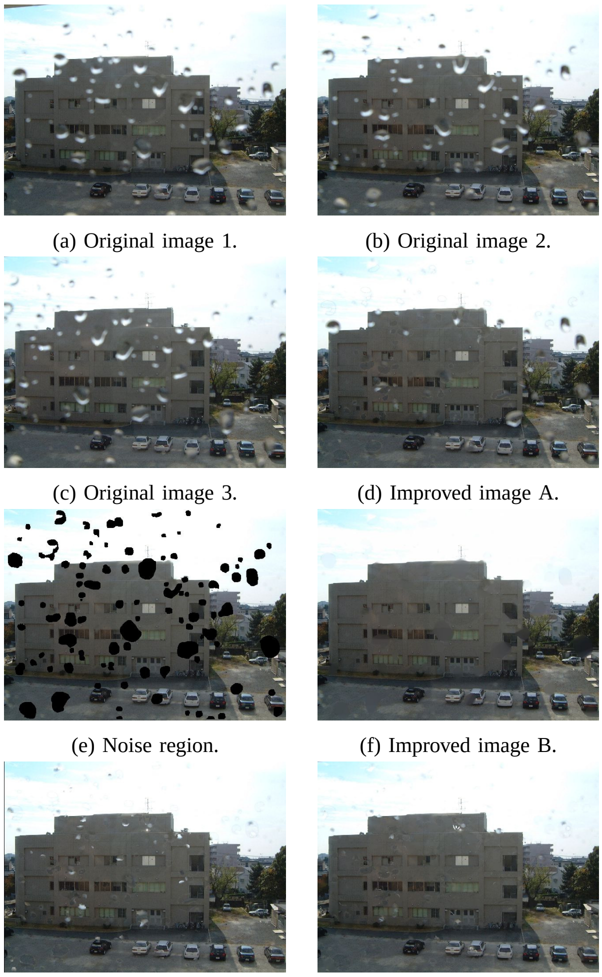
X. REFERENCES

[1] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald and P. J. W. Rayner: “Detection of Missing Data in Image Sequences,” *IEEE Transactions on Image Processing*, Vol.4, No.11, pp.1496–1508, 1995.

[2] A. Nagai, Y. Kuno and Y. Shirai: “Surveillance System Based on Spatio-Temporal Information,” *Proceedings of the 1996 IEEE International Conference on Image Processing*, Vol.2, pp.593–596, 1996.

[3] H. Hase, K. Miyake and M. Yoneda: “Real-time Snowfall Noise Elimination,” *Proceedings of the 1999 IEEE International Conference on Image Processing*, Vol.2, pp.406–409, 1999.

[4] T. Matsuyama, T. Ohya and H. Habe: “Background Subtraction for Non-Stationary Scene,” *Proceedings of the 4th Asian Conference on Computer Vision*, pp.662–667, 2000.



(a) Original image 1. (b) Original image 2.

(c) Original image 3. (d) Improved image A.

(e) Noise region. (f) Improved image B.

(g) Improved image C. (h) Improved image D.

Fig. 8. Experimental results IV.

[5] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald and P. J. W. Rayner: “Interpolation of Missing Data in Image Sequences,” *IEEE Transactions on Image Processing*, Vol.4, No.11, pp.1509–1519, 1995.

[6] S. Masnou and J.-M. Morel: “Level Lines Based Disocclusion,” *Proceedings of the 5th IEEE International Conference on Image Processing*, pp.259–263, 1998.

[7] L. Joyeux, O. Buisson, B. Besserer and S. Boukir: “Detection and Removal of Line Scratches in Motion Picture Films,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.548–553, 1999.

[8] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester: “Image Inpainting,” *Computer Graphics* (*Proceedings of SIGGRAPH2000*), pp.417–424, 2000.

[9] J. Weng, P. Cohen and M. 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.