The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Noises Removal from Image Sequences Acquired with Moving Camera by Estimating Camera Motion from Spatio-Temporal Information

Atsushi Yamashita, Isao Fukuchi and Toru Kaneko

Abstract— **This paper describes a method for removing adherent noises from image sequences. In outdoor environments, it is often the case that scenes taken by a camera are deteriorated because of adherent noises such as waterdrops on the surface of the lens-protecting glass of the camera. To solve this problem, our method takes advantage of image sequences captured with a moving camera whose motion is unknown. Our method estimates a camera motion only from image sequences, and makes a spatio-temporal image to extract the regions of adherent noises by examining differences of track slopes in cross section images between adherent noises and other objects. Finally, regions of noises are eliminated by replacing with image data corresponding to object regions. Experimental results show the effectiveness of our method.**

I. Introduction

In this paper, we propose a noise removal method from image sequences acquired with a moving camera whose motion is unknown by a spatio-temporal image processing. A spatio-temporal image can be generated by merging the acquired image sequence (Fig. 1(a)), and then cross-section images can be extracted from the spatio-temporal image (Fig. 1(b)). In these cross-section images, we can detect moving objects and estimate the motion of objects by tracing trajectories of edges or lines.

In recent years, cameras are widely used for surveillance systems in outdoor environments such as the traffic flow observation, the trespassers detection, and so on. It is also one of the fundamental sensors for outdoor robots. 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.

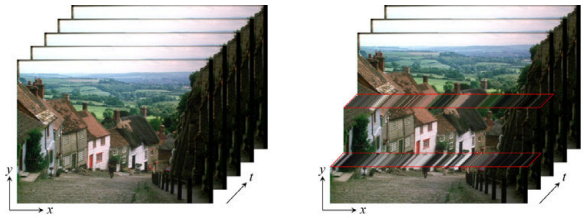
For example, waterdrops or mud blobs attached on the protecting glass may interrupt a field of view in rainy days. It would be desirable to remove adherent noises from images of such scenes for surveillance systems and outdoor robots.

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.

This research was partially supported by Special Project for Earthquake Disaster Mitigation in Urban Areas.

A. Yamashita, I. Fukuchi and T. Kaneko are with Department of Mechanical Engineering, Shizuoka University, 3–5–1 Johoku, Naka-ku, Hamamatsu-shi, Shizuoka 432–8561, Japan yamashita@ieee.org

I. Fukuchi is now with Brother Industries,Ltd., Japan



(a) Spatio-temporal image. (b) Cross-section.

Fig. 1. Spatio-temporal image.

Therefore, this paper proposes a new noise removal method from images by using image processing techniques.

A lot of image interpolation or restoration techniques for damaged and occluded images have been also proposed in image processing and computer vision societies [1]–[10]. However, some of them can only treat with line-shape scratches [1]–[3], 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) [4]–[10]. These methods are not suitable for surveillance systems and outdoor robots.

On the other hand, there are automatic methods that can remove noises without helps of human operators [11], [12]. Hase *et al.* have proposed a real-time snowfall noise elimination method from moving pictures by using a special image processing hardware [11]. Garg and Nayar have proposed an efficient algorithm for detecting and removing rain from videos based on a physics-based motion blur model that explains the photometry of rain [12]. These techniques work well under the assumptions that snow particles or raindrops are always falling. In other words, they can detect snow particles or raindrops because they move constantly.

However, adherent noises such as waterdrops on the surface of the lens-protecting glass may be stationary noises in the images. Therefore, it is difficult to apply these techniques to our problem because adherent noises that must be eliminated do not move in images.

To solve the static noise problem, we have proposed the method that can remove view-disturbing noises from images taken with multiple cameras [13], [15]. Previous study [13] is based on the comparison of images that are taken with multiple cameras. 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.

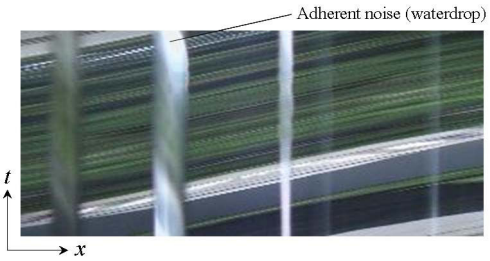


Fig. 2. Cross-section of spatio-temporal image (camera motion: rotation).

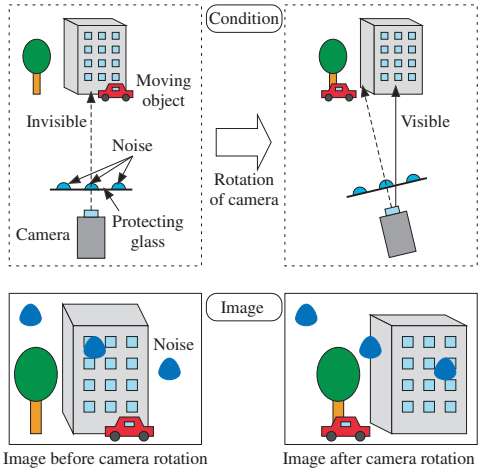


Fig. 3. Image acquisition by using camera rotation.

Therefore, we have proposed a method that can remove waterdrops from stereo image pairs that contain objects both in a distant scene and in a close range scene [14], [15]. This method utilizes the information of corresponding points between stereo image pairs, and thereby sometimes cannot work well when appearance of waterdrops differs from each other between left and right images.

We have also proposed a noise removal method by using a single camera [16]–[19]. These methods use a rotating camera, and eliminate adherent noises in the image sequence. These methods work well when the accurate camera motion is known. However, in real world application, it is often the case that the camera motion is unknown. In that case, it is necessary to estimate the direction and the velocity of the camera motion only from image sequences.

Therefore, in this paper, we estimate the camera motion from the information of spatio-temporal image sequences, and then detect and remove adherent noises in cross section images of the spatio-temporal image (Fig. 2).

II. Noise Detection and Removal Method

As to adherent noises on the protecting glasses of the camera, the positions of noises in images do not change when the direction of the camera changes (Fig. 3). This is because adherent noises are attached to the surface of the protecting glass of the camera and move together with the camera. On the other hand, the position of static background scenery and that of moving objects change while the camera rotates.

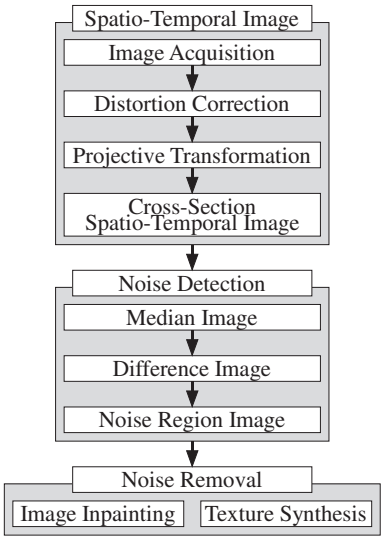


Fig. 4. Overview of our method.

We transform the image after the camera rotation to the image whose gaze direction (direction of the principal axis) is same with that before the camera rotation. Accordingly, we can obtain a new image in which only the positions of adherent noises and moving objects are different from the image before the camera rotates.

A spatio-temporal image is obtained by merging these transformed images. In the spatio-temporal image, trajectories of adherent noises can be calculated. Therefore, positions of noises can be also detected in the image sequence from the spatio-temporal image. Finally, we can obtain a noise-free image sequence by estimating textures on adherent noise regions.

The procedure of our method is shown in Fig. 4. The methods of noise removal and detection are based on [19].

III. Spatio-Temporal Image

A. Image Acquisition

An image sequence is acquired while a camera rotates .

At first (frame 0), one image is acquired where the camera is fixed. In the next step (frame 1), another image is taken after the camera rotates θ_1 rad about the axis which is perpendicular to the ground and passes along the center of the lens. In the t-th step (frame t), the camera rotate θ_t rad and the t-th image is taken. To repeat this procedure n times, we can acquire $n/30$ second movie if we use a 30fps camera.

B. Distortion Correction

The distortion from the lens aberration of images is rectified. Let (x', y') be the coordinate value without distortion, $(\tilde{x}, \tilde{y})$ be the coordinate value with distortion (observed coordinate value), and κ_1 be the parameter of the radial distortion, respectively [20]. The distortion of the image is corrected by (1) and (2).

$$\tilde{x} = x' + \kappa_1 x' \left(x'^2 + y'^2\right), \quad (1)$$

$$\tilde{y} = y' + \kappa_1 y' \left(x'^2 + y'^2\right). \quad (2)$$

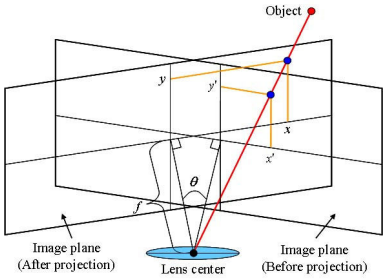


Fig. 5. Projective transformation.

C. Projective Transformation

In the next step, the acquired t-th image (the image after θ_t rad camera rotation) is transformed by using the projective transformation. The coordinate value after the transformation (x, y) is expressed as follows (Fig. 5):

$$x = f\frac{f\tan\theta_t + x'}{f - x'\tan\theta_t}, \quad (3)$$

$$y = f\frac{\sqrt{1+\tan^2\theta_t}}{f - x'\tan\theta_t}y', \quad (4)$$

where (x', y') is the coordinate value of the t-th image before transformation, and f is the image distance (the distance between the center of lens and the image plane), respectively.

The t-th image after the camera rotation is transformed to the image whose gaze direction is same with that before the camera rotation.

The estimation method of the direction and the angle of the camera rotation is explained in Chapter V.

D. Cross-Section of Spatio-Temporal Image

Spatio-temporal image $I(x, y, t)$ is obtained by arraying all the images in chronological order (Fig. 6(a)).

We can clip a cross-section image of $I(x, y, t)$. For example, Fig. 6(b) shows the cross-section image of the spatio-temporal image in Fig. 6(a) along $y = y_1$.

Here, let $S(x, t)$ be the cross-section spatio-temporal image. In this case, $S(u, t) = I(x, y_1, t)$.

In the cross-section spatio-temporal image $S(x, t)$, the trajectories of the static background scenery become vertical straight lines owing to the effect of the projective transformation. On the other hand, the trajectories of adherent noises in $S(x, t)$ become curves whose shapes can be calculated by (3) and (4)[1].

In this way, there is a difference between trajectories of static objects and those of adherent noises. This difference helps to detect noises.

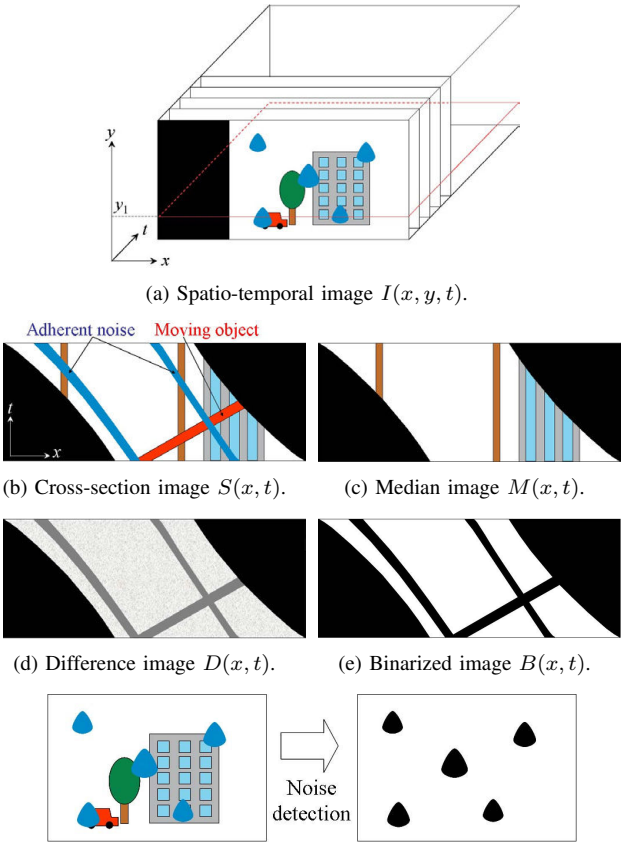
IV. Noise Detection

A. Median Image

Median values along time axis t are calculated in the cross-section spatio-temporal image $S(x, t)$. After that, a median image $M(x, t)$ can be generated by replacing the original pixel values by the median values (Fig. 6(c)).

Adherent noises are eliminated in $M(x, t)$, because these noises in $S(x, t)$ are small in area as compared to the background scenery.

A clear image sequence can be obtained from $M(x, t)$ by using the inverse transformation of (3) and (4) if there is no moving object in the original image. However, if the original image contains moving objects, the textures of these objects blur owing to the effect of the median filtering. Therefore, the regions of adherent noises are detected explicitly, and image restoration is executed for the noise regions to generate a clear image sequence around the moving objects.



(a) Spatio-temporal image $I(x, y, t)$.

(b) Cross-section image $S(x, t)$.

(c) Median image $M(x, t)$.

(d) Difference image $D(x, t)$.

(e) Binarized image $B(x, t)$.

(f) Noise region image $R(x, y)$.

Fig. 6. Spatio-temporal image processing.

B. Difference Image

A difference between the cross-section spatio-temporal monochrome image and the median monochrome image is calculated for obtaining the difference image $S(x, t)$ by (5).

$$D(x, t) = |S(x, t) - M(x, t)|. \quad (5)$$

Pixel values in regions of $D(x, t)$ where adherent noises exist become large, while pixel values of $D(x, t)$ in the background regions are small (Fig. 6(d)).

C. Noise Region Image

The regions where the pixel values of the difference images are larger than a certain threshold T_b are defined

[1]In Fig. 6(b), the trajectory of an adherent noise looks like a straight line, however, it is slightly-curved.

as the noise candidate regions. The binarized image $B(x,t)$ is obtained by

$$B(x,t) = \begin{cases} 0, & D(x,t) < T_b \\ 1, & D(x,t) \geq T_b \end{cases} . \quad (6)$$

The region of $B(x,t) = 1$ is defined as noise candidate regions (Fig. 6(e)). Note that an adherent noise does not exist on the same cross-section image when time t increases, because y-coordinate value of the adherent noise changes owing to the influence of the projective transformation in (4). Therefore, we consider the influence of this change and generate $B(x,t)$ in the way that the same adherent noise is on the same cross-section image.

In the next step, regions of adherent noises are detected by using $B(x,t)$. The trajectories of adherent noises are expressed by (3). Therefore, the trajectory of each curve is tracked and the number of pixel where $B(x,t)$ is equal to 1 is counted. If the total counted number is more than the threshold value T_n, this curve is regarded as the noise region. As mentioned above, this tracking procedure is executed in 3-D (x,y,t) space. This process can detect adherent noise regions precisely, even when there are moving objects in the original image sequence thanks to the probability voting (counting).

After detecting noise regions in all cross-section spatio-temporal image $S(x,t)$, the noise region image $R(x,y)$ is generated by the inverse projective transformation from all $B(x,t)$ information (Fig. 6(e)).

Ideally, the noise regions consist of adherent noises. However, the regions where adherent noises don't exist are extracted in this process because of other image noises. Therefore, the morphological operations (opening, *i.e.* erosion and dilation) are executed for eliminating small noises.

V. Estimation of Camera Motion

The direction and the angle of the camera rotation are estimated only from image sequences. At first, they are estimated by an optical flow. However, the optical flow may contain error. Therefore, the rotation angle is estimated between two adjacent frames by an exploratory way. Finally, the rotation angle is estimated between each frame and base frame.

A. Motion Estimation by Optical Flow

An optical flow between two images can be obtained by using the following equation [21].

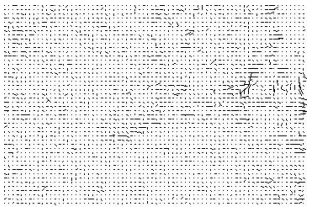
$$I_x u + I_y v + I_t = 0 \quad (7)$$

where $I(x,y,t)$ is an image sequence, optical flow $(u,v) = (dx/dt, dy/dt)$ is a velocity vector in the image, $(I_x, I_y) = (\partial I/\partial x, \partial I/\partial y)$ is a spatial gradient of pixel value, and $I_t = \partial I/\partial t$ is a temporal derivation of pixel value, respectively.

This equation contains two unknown parameters (u,v), and cannot be solved uniquely. We can solve the equation under the assumption that the optical flow in a certain local region is constant.

$$\mathbf{Ga} = -\mathbf{b}, \quad (8)$$

(a) Time t. (b) Time $t+1$.



(c) Optical flow by [21].

(d) Optical flow after step 1. (e) Optical flow after step 2.

Fig. 7. Optical flow.

where

$$\mathbf{G} = \begin{bmatrix} I_{1x} & I_{1y} \\ \vdots & \vdots \\ I_{ix} & I_{iy} \\ \vdots & \vdots \\ I_{N^2x} & I_{N^2y} \end{bmatrix}, \mathbf{a} = \begin{bmatrix} u \\ v \end{bmatrix}, \mathbf{b} = \begin{bmatrix} I_{1t} \\ \vdots \\ I_{it} \\ \vdots \\ I_{N^2t} \end{bmatrix} \quad (9)$$

The optical flow can be calculated by using the least square method.

$$\hat{\mathbf{a}} = -(\mathbf{G^T G})^{-1}\mathbf{G^T b}. \quad (10)$$

However, the obtained optical flow sometimes contains errors. In this paper, we reduce errors by considering the following two constraints.

Step 1: We can remove large optical flow that exceeds the maximum velocity of the camera (u_{max}, v_{max}). In addition, small optical flow that is less than (u_{min}, v_{min}) can be removed.

Step 2: The rotation direction of the camera in each frame is whether clockwise or counterclockwise. The rotation direction can be estimated by calculating the average optical flow of whole image plane after Step 1. Therefore, counterrotating optical flow against the average flow can be removed.

Fig. 7 shows the calculation result of the optical flow. Removal results in Step 1 are indicated as blue lines, and those in step 2 are indicated as red lines, respectively.

After Step 1 and 2, the average optical flow (u_{ave}, v_{ave}) in the whole image plane is calculated. The rotation angle

can be estimated by (3) (Initial estimation angle θ_{start}).

$$\theta_{start} = \tan^{-1} \frac{u_{ave}}{f}, \tag{11}$$

However, θ_{start} may still contain errors. Then, we introduce the following estimation method.

B. Motion Estimation between Two Adjacent Frames

Estimation of the rotation angle between two adjacent frame can be achieved by considering the difference between $I(x, y, t)$ and $I(x, y, t + 1)$ after the projective transformation.

At first, average brightness of $I(x, y, t)$ and $I(x, y, t + 1)$ are normalized. Then, Euclidean distance between $I(x, y, t)$ and $I(x, y, t + 1)$ in the RGB color space ($D_{rgb}(x, y)$) is calculated. If the estimated rotation angle is accurate, the difference between $I(x, y, t)$ and $I(x, y, t + 1)$ becomes small. Therefore, the sum of $D_{rgb}(x, y)$ in all common field of view becomes small in the case of accurate rotation angle.

$$D_{ave} = \frac{1}{M} \sum_x \sum_y D_{rgb}(x, y), \tag{12}$$

where M is a number of pixel in common field of view between two frames.

Then, we can search θ_{opt1} in greedy search algorithm by solving the following equation.

$$\theta_{opt1} = \arg\min_{\theta} D_{ave}. \tag{13}$$

C. Motion Estimation between Each Frame and Base Frame

After the estimation method mentioned in Section V-B, we can obtain the almost exact angle. However, accumulated error increases when $I(x, y, t)$ is transformed according to the base frame image $I(x, y, t_{base})$. Finally, the optimal rotation angle $\theta_{opt2} (= \theta_t)$ is calculated. In this case, initial estimation angle is calculated as follows.

$$\theta'_{start} = \sum_{t_{base}}^{t} \theta_{opt1}. \tag{14}$$

Then, we search the optimal rotation angle θ_{opt2} from θ'_{start} in a similar way of Section V-B.

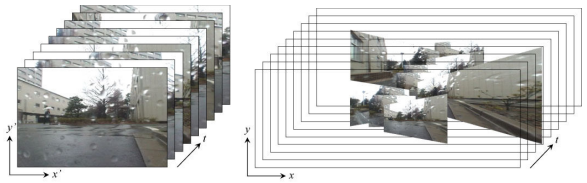
VI. Noise Removal

Adherent noises are eliminated from the cross-section spatio-temporal image $S(x, t)$ by using the image restoration technique [7] for the noise regions detected in Section IV.

At first, an original image $S(x, t)$ is decomposed into a structure image $f(x, t)$ and a texture image $g(x, t)$ [22].

After the decomposition, the image inpainting algorithm [4] is applied for the noise regions of the structure image $f(x, t)$, and the texture synthesis algorithm [23] is applied for the noise regions of the texture image $g(x, t)$, respectively. This method [7] overcomes the weak point that the original image inpainting technique [4] has the poor reproducibility for a complicated texture. After that, noise-free image can be obtained by merging two images.

Finally, a clear image sequence without adherent noises is created with the inverse projective transformation.



(a) $I(x', y', t)$. (b) $I(x, y, t)$.

Fig. 8. Spatio-temporal image (waterdrop).

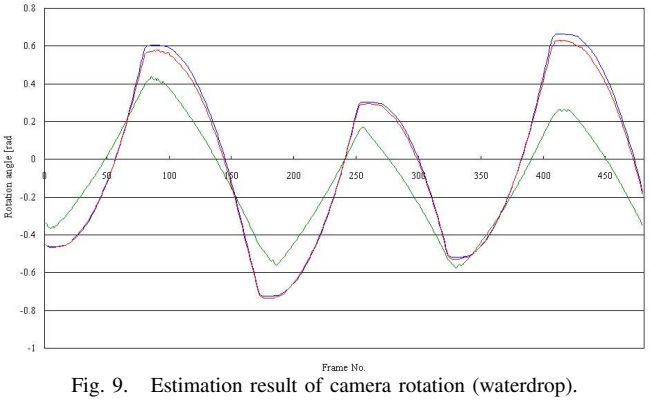


Fig. 9. Estimation result of camera rotation (waterdrop).

VII. Experiment

Image sequence was acquired in the outdoor environment. We used a pan-tilt-zoom camera (Sony EVI-D100) whose image distance f was calibrated as 261pixel.

Fig. 8(a) shows an example of the original spatio-temporal image, and Fig. 8(b) shows the result of the projective transformation, respectively.

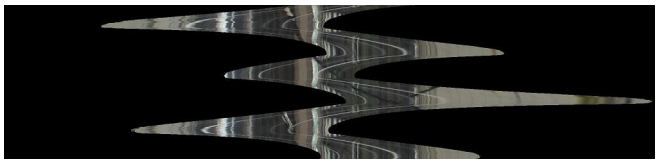
Fig. 9 shows the estimation result of camera rotation angle. In Fig. 9, green line indicates the estimation result from optical flow (Section V-A), blue one indicates that of Section V-B, and red one indicates the final result (Section V-C), respectively. The final estimation result coincides with the ground truth very well (maximum error: within 0.1deg), although the result of optical flow has error (maximum error: 0.4deg).

Figs. 10 and 11 show the intermediate results of the noise removal. Figs. 10(a) and (b) show the cross-section spatio-temporal image $S(x, t)$ in color and monochromic formats, respectively. There is a moving object (a human with a red umbrella) in this image sequence. Figs. 11(a), (b), (c) show the median image $M(x, t)$, the difference image $D(x, t)$, and the binarized image $B(x, t)$, respectively. Fig. 11 (d) shows the noise region image $R(x, y)$, and Fig. 11 (e) shows the noise removal result from the cross-section spatio-temporal image, respectively.

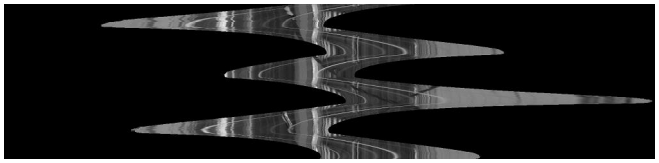
Fig. 12 shows the final results of noise removal for the image sequence. All waterdrops are eliminated and the moving object can be seen very clearly in all frames.

Figs. 13 and 14 show results of mud blob removal.

From these results, it is verified that our method can remove adherent noises on the protecting glass of the camera regardless of their positions, colors, sizes, existence of mov-



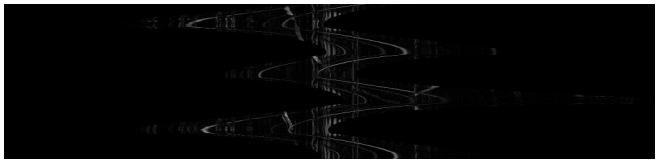
(a) Color.



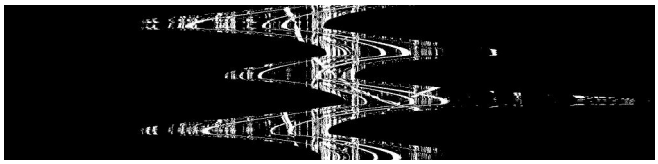
(b) Grayscale.

Fig. 10. Cross-section spatio-temporal image $S(x,t)$ (waterdrop).

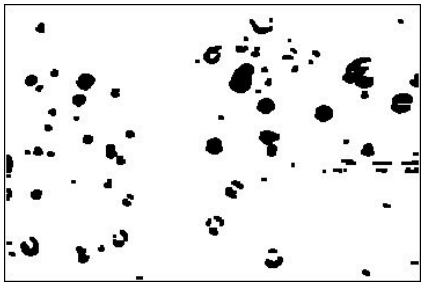
(a) Median image $M(x,t)$.



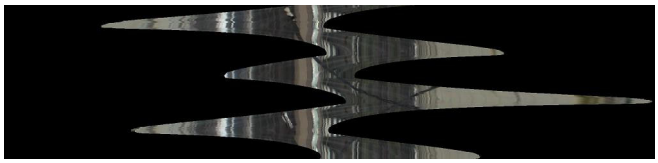
(b) Difference image $D(x,t)$.



(c) Binarized image $B(x,t)$.



(d) Noise region image $R(x,t)$.



(e) Cross-section spatio-temporal image after noise removal.

Fig. 11. Result of noise detection (waterdrop).

ing objects, and the direction and the speed of the camera rotation.

Fig. 15 shows comparison results of the rotation angle estimation. Fig. 15(a) shows the result of Section V-A, Fig. 15(b) shows the result of Section V-B, Fig. 15(c) shows the result of Section V-C, respectively. In Figs. 15(a) and (b), trajectories of objects are not parallel to the time axis, while the results of our method (Figs. 15(c)) are almost parallel to time axis.

To verify the accuracy of the noise detection, the results of our methods (Figs. 17(c) and 17(c)) are compared with the ground truth that is generated by a human operator manually (Figs. 17(b) and 17(b)). In Figs. 16(d) and 17(b), red regions indicate the correct detection, blue regions mean undetected noises, and green regions are false (exceeded) detection regions. Actually, undetected noises are hard to detect when we see the final result (Fig. 12(b)). This is because the image interpolation works well in the noise removal step.

Figs. 18 and 19 show comparison results of texture interpolation with an existing method. In these figures, (a) shows the ground truth, (b) shows the position of noise region, (c) shows the noise removal result, and (d) shows the difference between (a) and (c). Fig. 18(c) shows the result by the image inpainting technique [7], and Fig. 19(c) shows the result by our method. The result by the existing method is not good (Fig. 18(d)), because texture of the noise region is estimated only from adjacent region. In principle, it is difficult to estimate texture in several cases from only a single image. On the other hand, our method can estimate texture robustly by using a spatio-temporal image processing (Fig. 19(d)).

From these results, the validity of our method is verified.

VIII. Conclusion

In this paper, we propose a noise removal method from image sequence acquired with a rotating camera whose motion is unknown. The camera motion is estimated by using an optical flow and an additional optimization method. We makes a spatio-temporal image to extract the regions of adherent noises by examining differences of track slopes in cross section images between adherent noises and other objects. Regions of adherent noises are interpolated from the spatio-temporal image data.

Experimental results show the effectiveness of our method even when the camera motion is unknown.

As future works, a camera translation should be considered in addition to a camera rotation.

References

[1] 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.

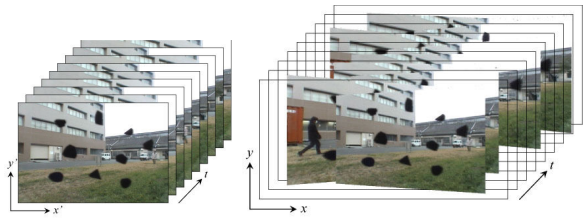
[2] S. Masnou and J.-M. Morel: "Level Lines Based Disocclusion," *Proceedings of the 5th IEEE International Conference on Image Processing* (*ICIP1998*), pp.259–263, 1998.

[3] 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.

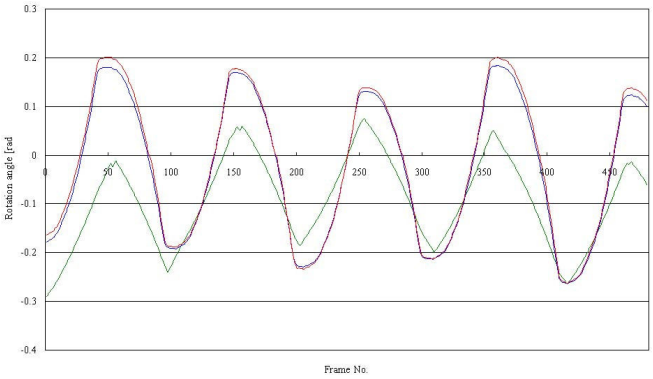
[4] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester: "Image Inpainting," *ACM Transactions on Computer Graphics* (*Proceedings of SIGGRAPH2000*), pp.417–424, 2000.

(a) Original image. (b) Result image.

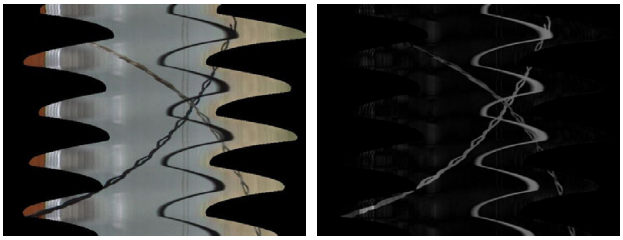
Fig. 12. Result of noise removal (waterdrop).



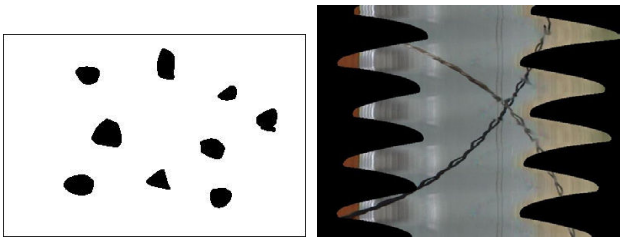
(a) $I(x', y', t)$. (b) $I(x, y, t)$.



(c) Estimation result of camera rotation.



(d) $S(x, t)$. (e) $D(x, t)$.



(f) $R(x, t)$. (g) Noise removal.

Fig. 13. Result of noise detection (mud blob).

[5] M. Bertalmio, A. L. Bertozzi and G. Sapiro: "Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (*CVPR2001*), Vol.1, pp.355–362, 2001.

[6] S. H. Kang, T. F. Chan and S. Soatto: "Inpainting from Multiple Views," *Proceedings of the 1st International Symposium on 3D Data Processing Visualization and Transmission*, pp.622–625, 2002.

[7] M. Bertalmio, L. Vese, G. Sapiro and S. Osher: "Simultaneous Structure and Texture Image Inpainting," *IEEE Transactions on Image Processing*, Vol.12, No.8, pp.882–889, 2003.

[8] Y. Matsushita, E. Ofek, X. Tang and H.-Y. Shum: "Full-frame Video Stabilization," *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (*CVPR2005*), Vol.1, pp.50–57, 2005.

[9] Y. Shen, F. Lu, X. Cao and H. Foroosh: "Video Completion for Perspective Camera Under Constrained Motion, *Proceedings of the 18th International Conference on Pattern Recognition* (*ICPR2006*), Vol.3, pp.63–66, 2006.

[10] Y. Wexler, E. Shechtman and M. Irani: "Space-Time Completion of Video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.29, No.3, pp.463–476, 2007.

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

[12] K. Garg, S. 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.

[13] A. Yamashita, M. Kuramoto, T. Kaneko and K. 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.

[14] A. Yamashita, Y. Tanaka and T. Kaneko: "Removal of Adherent Waterdrops from Images Acquired with Stereo Camera," *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS2005*), pp.953–958, 2005.

[15] Y. Tanaka, A. Yamashita, T. Kaneko and K. T. Miura: "Removal of Adherent Waterdrops from Images Acquired with a Stereo Camera System," *IEICE Transactions on Information and Systems*, Vol.89-D, No.7, pp.2021–2027, 2006.

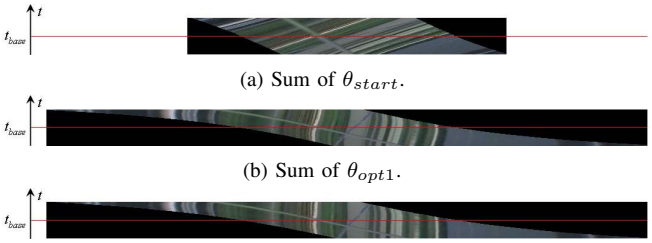
[16] A. Yamashita, T. Kaneko and K. T. Miura: "A Virtual Wiper – Restoration of Deteriorated Images by Using a Pan-Tilt Camera–," *Proceedings of the 2004 IEEE International Conference on Robotics and Automation* (*ICRA2004*), pp.4724–4729, 2004.

[17] A. Yamashita, T. Harada, T. Kaneko and K. 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.

[18] A. Yamashita, T. Harada, T. Kaneko and K. 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–

(a) Original image. (b) Result image.

Fig. 14. Result of noise removal (mud blob).



(a) Sum of θ_{start}.

(b) Sum of θ_{opt1}.

(c) Sum of θ_{opt2}.

Fig. 15. Accuracy of rotation angle estimation.

310, 2005.

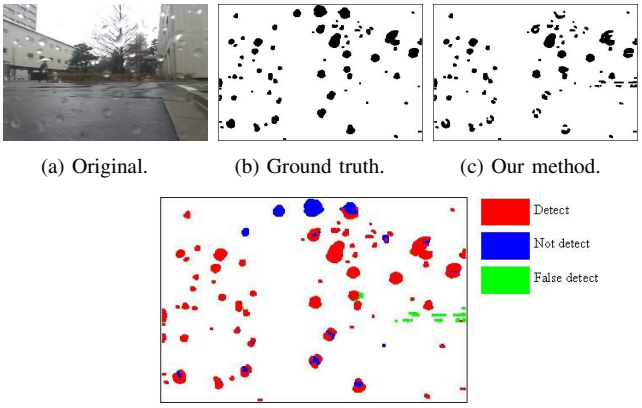
[19] A. Yamashita, I. Fukuchi, T. Kaneko and K. T. Miura: "Removal of Adherent Noises from Image Sequences by Spatio-Temporal Image Processing," *Proceedings of the 2008 IEEE International Conference on Robotics and Automation* (*ICRA2008*), pp.2386–2391, 2008.

[20] 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.

[21] B. D. Lucas and T. Kanade: "An Iterative Image Registration Technique with an Application to Stereo Vision," *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pp.674–679, 1981.

[22] L. I. Rudin, S. Osher and E. Fatemi: "Nonlinear Total Variation Based Noise Removal Algorithms," *Physica D*, Vol.60, pp.259–268, 1992.

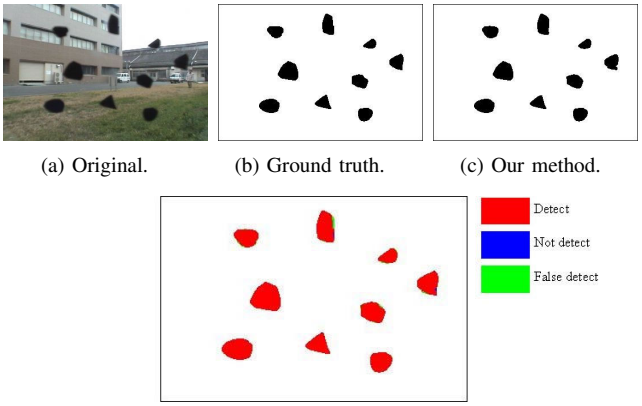
[23] A. A. Efros and T. K. Leung: "Texture Synthesis by Non-parametric Sampling," *Proceedings of the 7th IEEE International Conference on Computer Vision* (*ICCV1999*), Vol.2, pp.1033–1038, 1999.



(a) Original. (b) Ground truth. (c) Our method.

(d) Comparison.

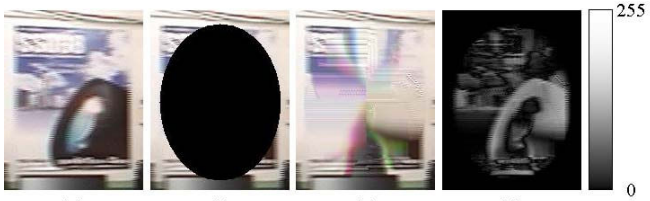
Fig. 16. Accuracy of noise removal (waterdrop).



(a) Original. (b) Ground truth. (c) Our method.

(d) Comparison.

Fig. 17. Accuracy of noise removal (mud blob).



(a) (b) (c) (d)

Fig. 18. Accuracy of noise removal (one frame [7]).

(a) (b) (c) (d)

Fig. 19. Accuracy of noise removal (our method).