Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Video Monitoring of Slope Failure Using Spatiotemporal Gabor Filtering

Ken Okamoto, Toshio Watanabe, Hiroshi Ban, Yuji Maeda
NTT Energy and Environment Systems Laboratories
Nippon Telegraph and Telephone Corporation
Tokyo, Japan
okamoto.ken@lab.ntt.co.jp

Akitoshi Hanazawa, Takashi Morie
Graduate School of Life Science and Systems Engineering
Kyushu Institute of Technology
Fukuoka, Japan

***Abstract*—We propose a method for detecting precursors, such as small rock and/or soil fall, which occur prior to massive slope failure. The key feature of our method is directly recognizing the trajectory of a small collapse using spatiotemporal Gabor filtering. Simulation analysis, where the conditions of the simulation are quantitatively defined, reveals the effectiveness of the proposed method in detecting a tiny moving object with low contrast in the background under low frame-rate video monitoring. Experiments using actual monitoring videos of a hazardous slope confirmed the effectiveness of our method. The effects of error factors in an outdoor environment, which may inhibit recognition, are also evaluated.**

***Keywords*—spatiotemporal Gabor filtering, video monitoring, tiny moving object, low contrast, low frame-rate, slope failure, precursors, trajectory**

I. INTRODUCTION

Many slope failures occur worldwide every year and sometimes cause significant damage to human life [1][2]. Even after the disaster, the slope remains in danger of failure due to the unstable ground of the slope. Therefore, rescue, survey, and recovery operations are faced with the hazards of secondary disasters [3]. Small falling rocks and/or soil failure (hereinafter called 'small collapse') usually occur a few hours prior to a massive failure. Although such small collapses are usually observed by human monitoring on site, human monitoring of a wide area of the slope is not easy, and support by machine vision technologies is necessary.

A small collapse is a type of a moving object. Although few studies deal with recognition of small collapses, many studies on recognizing moving objects, such as automobiles and humans, have been reported [4][5]. Because video images generally contain flicker and compression noise, discrimination of the target from these types of noise is necessary. In recognizing automobiles or humans, the motion and size of the target in a video scene are much different from the noise, making discrimination easy. However, due to peculiar difficulties, such as the target smallness on video images and its indistinguishable color from the background (slope surface), conventional approaches seem unsuitable for recognizing small collapses.

We propose a method for extracting small collapses using spatiotemporal Gabor filtering (STGF), which is effective in representing a model of low-level processing in visual perception [6][7]. The proposed method discriminates small collapses from noise in video images by focusing on the trajectory in the spatiotemporal space made by the small collapse. We show the basic characteristics of our method confirmed using simulation, where a quantitative artificial small collapse was superimposed on an actual slope image in a hazardous area. We also show the results where the proposed method was applied to a monitoring video of an actual hazardous area and discuss the practical factors in an outdoor environment, which may inhibit recognition.

II. METHOD USING SPATIOTEMPORAL GABOL FILTERING

A. Overview

Figure 1 outlines the procedure of the proposed method, consisting of three processes: preprocessing, feature extraction and recognition. The key feature of our method is the detection of a visually unclear small collapse by motion energy calculated using STGF in the feature extraction process. A

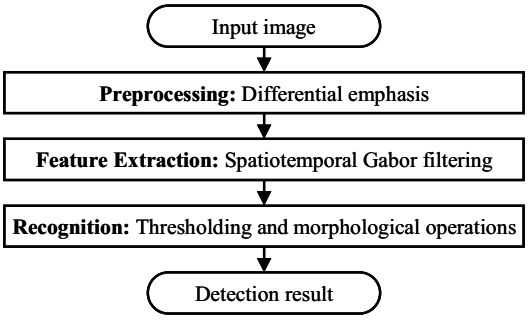

Figure 1. Procedure of proposed method.

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

trajectory of a small collapse in a short time forms a straight line in spatiotemporal space, while flicker noise exists as an isolated point. This difference is used to distinguish the small collapses from noise. Concretely, the trajectory of a small collapse is separated from the noise and extracted by the motion energy obtained using STGF. Details of STGF are described in the next section.

In preprocessing, since the intensity differential caused by the small collapse is very small, it is emphasized by the differentials of consecutive video frames in the following process:

$$I'_t = I_t + \alpha \left| I_t - I_{t-1} \right|^n , \tag{1}$$

where I_t is the intensity of an input image at time t, I'_t is the intensity after differential emphasis, α is a weight factor, and n is an exponential factor. Although noise is emphasized as well, STGF in the feature extraction process is well tolerant of noise. We empirically set $\alpha = 0.3$ and $n = 2.0$.

The recognition process identifies small collapses with basic image processing: thresholding and morphological operations, because STGF separates well the trajectory from isolated noises.

B. Spatiotemporal Gabor Filtering

STGF consists of a convolution operation on the input image I and a spatiotemporal Gabor kernel g_φ with arbitrary phase φ and the motion energy E is calculated as follows:

$$r_\varphi = I * g_\varphi , \tag{2}$$

$$E = \sqrt{r_0^{\,2} + r_{\pi/2}^{\,2}} , \tag{3}$$

where r is the response intensity. Since E is the square-root of the sum of the squares of the two response intensities, where their phases differ by $\pi/2$ from each other, the motion energy becomes phase-independent.

The spatiotemporal Gabor kernel is a cosine wave localized with the two-dimensional (2D) Gaussian function. The kernel at a point (x_0, y_0, t_0) in the spatiotemporal space is described by the following equations:

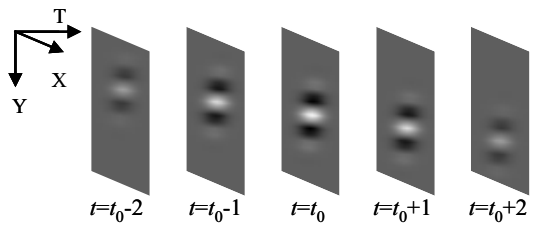

Figure 2. Example of spatiotemporal Gabor kernel where spread of Gaussian window in time direction σ_t is 2 (frames). Bright region has positive value, dark region has negative value, and gray region has zero value.

$$g_\varphi = \cos\left(\frac{2\pi}{\lambda}\left(\bar{x} + v\bar{t}\right) + \varphi \right) \times \frac{1}{\sqrt{2\pi}^3 \sigma_x \sigma_y \sigma_t} \exp\left(-\left(\frac{\left(\bar{x} + v\bar{t}\right)^2}{2\sigma_x^2} + \frac{\bar{y}^2}{2\sigma_y^2} + \frac{\bar{t}^2}{2\sigma_t^2} \right) \right), \tag{4}$$

$$\bar{x} = \left(x - x_0\right)\cos\left(\theta\right) + \left(y - y_0\right)\sin\left(\theta\right), \tag{5}$$

$$\bar{y} = -\left(x - x_0\right)\sin\left(\theta\right) + \left(y - y_0\right)\cos\left(\theta\right), \tag{6}$$

$$\bar{t} = t - t_0 , \tag{7}$$

where λ is the wavelength, v and θ are the velocity and the moving direction in the X-Y plane at the center point of the Gaussian function, respectively, and σ_x, σ_y, σ_t is the spread of the Gaussian window in each axis. Here, the kernel size in the X-Y plane is defined as 2λ. Figure 2 shows an example of spatiotemporal Gabor kernel detecting a downward motion.

STGF indicates a strong response for a moving object, where the velocity and the moving direction of the object correspond to parameters v and θ of STGF, respectively. By using various sets of parameters θ and v and choosing the set that shows maximum motion energy, the velocity and the moving direction at each pixel can be identified.

III. Experiments Using Simulation

A. Quantitative study using simulation

We examined the characteristics of the proposed method by simulation using an artificial small collapse superimposed on an actual video of a steep slope where a massive landslide occurred (Fig. 3). The artificial small collapse was a 4×4 square bitmap under the quantitatively defined conditions listed in Table I. Each pixel color was set randomly with a Gaussian distribution and was reset frame by frame. An example of this artificial small collapse is shown in Fig. 4.

TABLE I. Parameters of Small Collapse

Parameter (Unit)	Setup
Size (pixel)	4 × 4
Mean intensity difference from background (8-bit scale)	−5
Intensity distribution (8-bit scale) (Standard deviation of pixels)	2
Falling speed (pixel / frame)	2
Moving direction (rad)	Downward (−π/2)

TABLE II. Parameters of Spatiotemporal Gabor Kernel

Parameter (Unit)	Setup
Wave length: λ (pixel)	4
Detection Range of velosity: v (pixel / frame)	−10 to 10
Moving direction: θ (rad)	−π/2 (downward)

The feature of the proposed method is that STGF well discriminates a small collapse from background noise in the feature extraction process. We evaluated the performance of STGF by using a receiver-operating characteristic (ROC) analysis, where an area with large motion energy was assumed as a collapse and an area under ROC curve (AUC) was used as a performance index. In the evaluation, the inside of the polygon shown in Fig. 3 was used as the monitoring area, excluding the river.

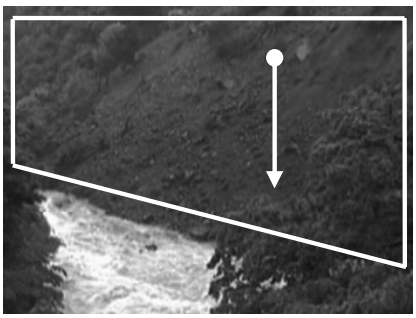

Figure 3. Simulation video. Artificial small collapse moves along arrow. Area enclosed by line is monitoring area. Background video was provided by Etsumi Mountain Area Sabo Office [8].

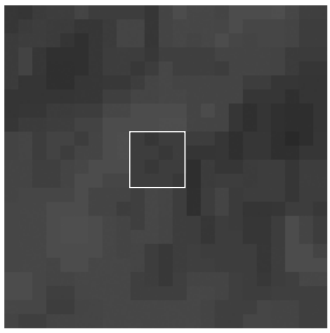

Figure 4. Zoomed image of artificial small collapse superimposed on actual image. Area enclosed in 4×4 pixel white box is small collapse.

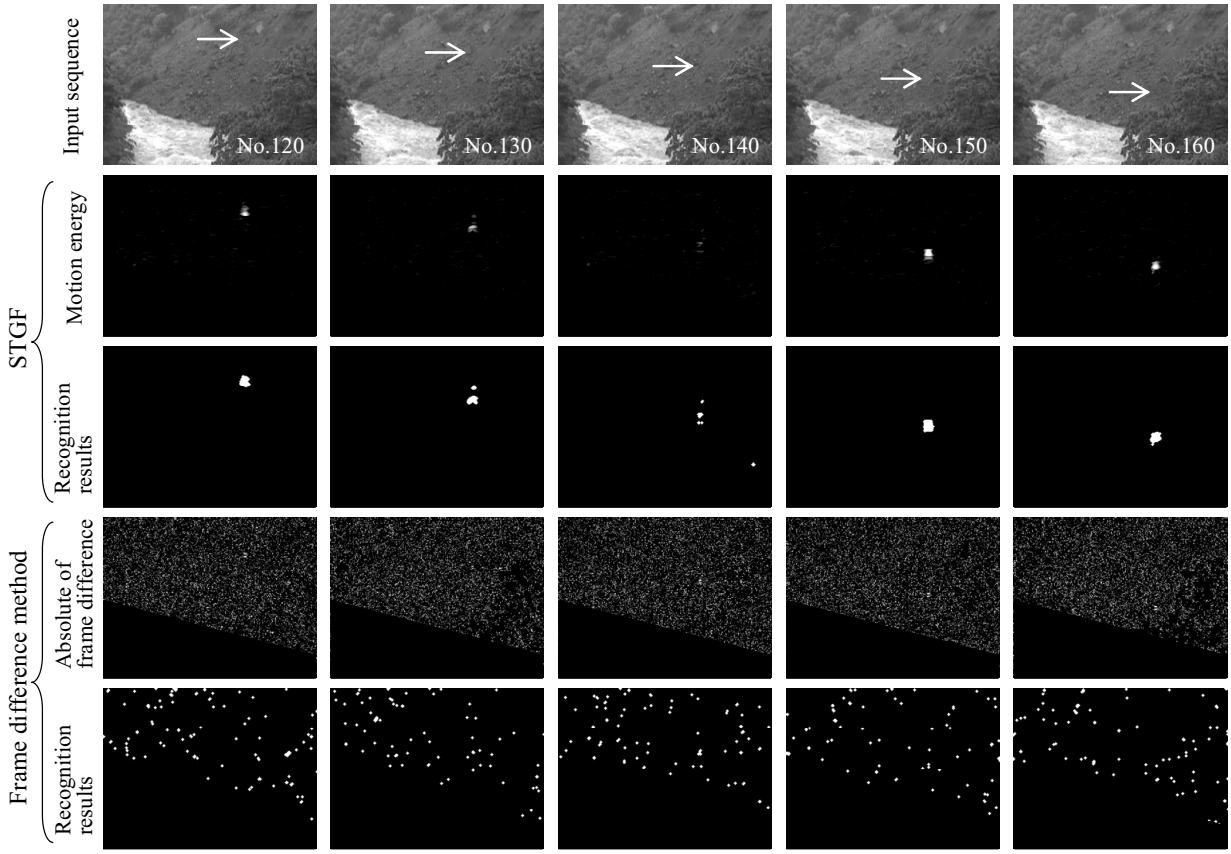

Figure 5. Comparison between proposed method and frame difference method. Arrow points to position of small collapse. Second and fourth rows are series of motion enegy of STGF and absolute of frame difference, respectively. Third and fifth rows are results of recognition based on second and fourth rows, respectively.

B. Comparison with Frame Difference Method

We evaluated the capability and effectiveness of STGF by comparing it with the frame difference method, which is a basic motion detection technique. Figure 5 shows the results of the comparison. The second row shows motion energy distribution obtained using STGF of which parameters are listed in Table II. The small collapse was discriminated from noise and was successfully tracked using STGF followed by thresholding and morphological operations, which are basic image processes, as shown in the third row in Fig. 5. However, in the frame difference method, the small collapse and the background noise were indistinguishable and the small collapse could not be captured with thresholding and morphological operations, as shown in the fourth and fifth rows. In both cases, the thresholds were determined by a discriminant analysis method and the morphological operations consisted of one erosion operation and two dilation operations with a 3×3 structure element.

The probability distributions of the motion energy obtained using STGF and the absolute of frame difference are shown in Fig. 6. While the distribution peaks of the small collapse are larger than those of the background in both cases, the results from STGF have a clearer distribution difference. The results of ROC analysis are shown in Fig. 7. The ROC curve of STGF with a more upper-left convex shape indicates that STGF has a better discrimination performance. The AUC scores of STGF and the frame difference method were 0.87 and 0.72, respectively. If a low hit-rate is allowed, the frame difference method is also enough. However, STGF is more effective when a high hit-rate is required.

We also evaluated the recognition performance by using recall, R, and precision, P, defined by the following equations:

$$R = \text{true positive} / (\text{true positive} + \text{false negative}), \quad (8)$$

$$P = \text{true positive} / (\text{true positive} + \text{false positive}). \quad (9)$$

Figure 8 shows the recognition performance with various thresholds. STGF has significantly better results compared with the frame difference method. These results correspond to the results shown in Fig. 5.

Although introducing advanced information processing techniques may improve the results of the frame difference method, the results showed the effectiveness of STGF on feature extraction. Having high recognition performance of STGF with a simple recognition process is robust against various disturbing factors of video monitoring in a natural environment.

C. Basic Characteristics of Proposed method

We examined the basic characteristics of the proposed method concerning the size of a small collapse, the contrast of the collapse to the background, and decreasing frame rate of the video image.

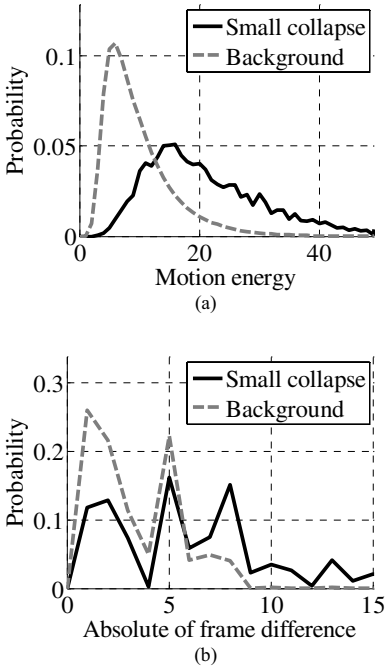

Figure 6. Comparison of probability density distribution between proposed method (a) and frame difference method (b). Region of small collapse (solid line) consists of all pixels in small collapse positions in all frames and has no margin around small collapse. Region of background (dashed line) consists of all pixels except those in small collapse positions in monitoring area.

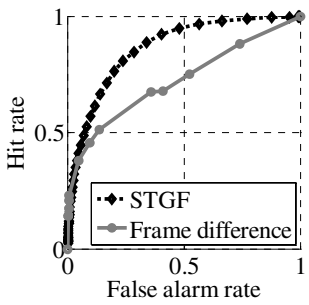

Figure 7. Comparison of feature extraction performance between STGF and frame difference method in ROC analysis.

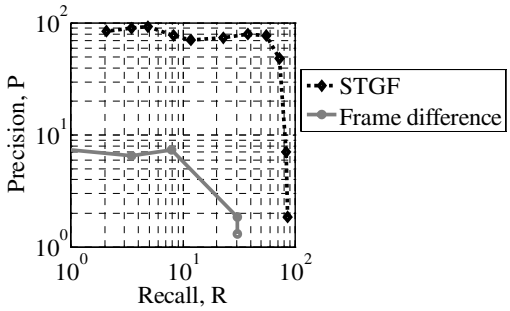

Figure 8. Discriminant analysis result from comparing recognition performance between STGF and frame difference method.

1) Dependence on Size of Collapse Region and STFG Wavelength

In general, there should be an appropriate value in wavelength λ depending on the target object size. Figure 9 shows AUC scores dependent on the size of a small collapse and STGF wavelength λ. Contrary to the prior expectations, when $\lambda = 4$, STGF marked the best AUC score in all collapse sizes. According to our observation, motion energy distribution was more affected by the texture of the background slope when the wavelength λ was larger. We will continue to investigate the relationship between the size of the collapse and STGF wavelength.

2) Dependence on Contrast between Small Collapse and Background

The dependence of discrimination performance on the intensity difference between a small collapse and the background and the strength of noise was examined. The mean intensity difference from the background in Table I varied from -20 to 20 (in 8-bit scale). Gaussian noise with standard deviation σ_n (= 0, 5, 10) was also added to the simulation image (Table III).

The results of the ROC analysis are shown in Fig. 10. Basically, higher contrast and/or less noise lead to a higher separation performance. However, noise distribution of up to 5 (in 8-bit scale) did not affect the discrimination performance.

3) Dependence on Frame Rate

The relation between the feature extraction performance and the frame rate of input video is shown in Fig. 11. AUC scores were calculated when the frame rate was 2, 3, 5, 10, 15, and 30 fps. Videos with lower frame-rate were made from the original simulation video of 30 fps. When the frame rate was more than 10 fps, the performance did not lower. This stable property ensures the robustness of the proposed method under low frame-rate.

At operation sites in hazardous areas, stable communication networks are not always available. The above feature is an advantage in practical applications. Our software used in the experiment only works for falling speeds of up to 10 pixels/frame, which means a frame rate higher than 6 fps. Therefore, the score dramatically declined when the frame rate was 2, 3, and 5 fps.

IV. Application to Actual Natural Scenes

We applied the proposed method to actual monitoring videos of a hazardous slope in an area heavily damaged by the Iwate-Miyagi Nairiku Earthquake in 2008, whose estimated magnitude was 7.2. Figure 12 shows an example of the steep slope monitored under recovery operation. The quadrangles in the figure indicate areas estimated where small collapses occurred. We detected 13 of 20 falling rocks and 7 of 18 slope failures in the monitoring video with our method. This detection rate of about 50% may seem too low. However, only with careful and repeated human video-checking, all small collapses were detected because some of these collapses were difficult to visually recognize with only one check. Therefore, this result implied the potential of the proposed method.

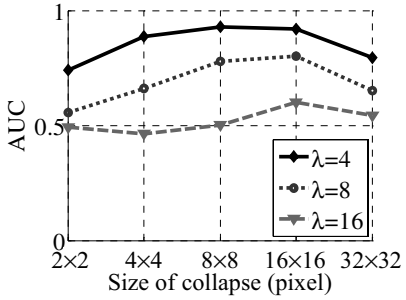

Figure 9. Discrimination performance dependent on collapse-region size and STGF wavelength λ.

TABLE III. Additional Noise and Image Quality

Noise level (σ_n)	Impression
0	No additional noise (original quality)
5	Noise is almost invisible
10	Clearly visible, but not so disruptive

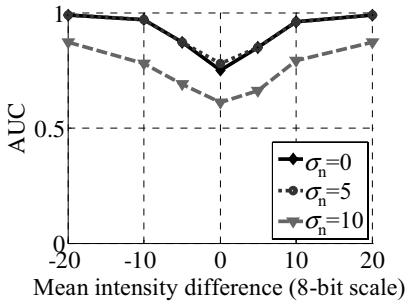

Figure 10. Discrimination performance dependent on mean intensity difference and background noise σ_n.

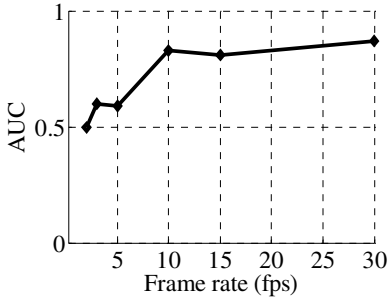

Figure 11. Discrimination performance dependent on frame rate.

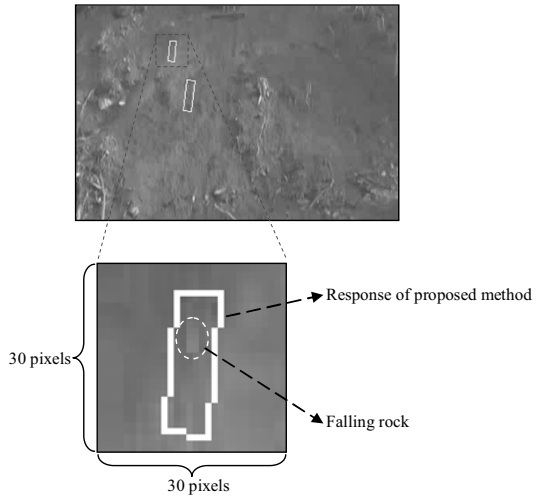

Figure 12. Example of recognition result of actual hazardous slope using proposed method: overview and zoomed image.

Figure 13. Monitoring in foggy weather.

Influences by natural disturbances on monitoring, such as fog (Fig. 13), motion of trees, and rainfall, were also examined. Fog did not cause false detection; however, since fog decreases the contrast of images, the performance of the recognition might decline. To tree movement or heavy rain, where raindrops were visible, STGF was rather sensitive. For discriminating between small collapses and such motions, time-series of motion and/or differences in falling speed should be considered.

V. CONCLUSION

We proposed a method for detecting precursors of slope failure, which are difficult to visually recognize due to its smallness and similar color with the background. Our method is used to directly detect trajectories of small collapses using STGF. The simulation results and the practical examinations using actual videos showed that the proposed method is effective in detecting tiny moving objects with low contrast to the background. The robustness to noise and frame-rate decline of the method is an advantage in practical monitoring applications.

Although there are various sensors for slope failure or landslide monitoring, video monitoring has the unique advantage that various events in a wide area can be observed simultaneously. Introducing image processing technologies enhances the capabilities of disaster monitoring and prevention systems. We will conduct more practical experiments and try to improve and enhance the proposed method.

ACKNOWLEDGMENT

We thank the Ministry of Land, Infrastructure, Transport and Tourism, Etsumi Mountain Area Sabo Office for permitting the use of their video.

REFERENCES

[1] F. Guzzetti, "Landslide fatalities and the evaluation of landslide risk in Italy," Engineering Geology, Vol. 58, No. 2, pp. 89-107, 2000.

[2] S. Cari and P. Scott H., "An investigation into the role of slope stability on archaelogocal site distibution and preservation in the middle sangro valley," The Geological Society of America annual meeting, 2006.

[3] M. Nagai, S. Kuroda, Y. Yuuki, T. Chen, and R. Shibasaki, "Surveying of the landslides in the mid Niigata prefecture earthquake by an unmanned helicopter," Asian Conference on Remote Sensing, 2005.

[4] C. Stauffer, and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 747-757, 2000.

[5] R. T. Collins, A. J. Lipton, H. Fujiyoshi, and T. Kanade, "Algorithms for cooperative multisensor surveillance," Proceedings of the IEEE, vol. 89, no. 10, pp. 1456-1477, 2001.

[6] E. H. Adelson, and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," J. Opt. Soc. Am. A, vol. 2, no. 2, pp. 284-299, 1985.

[7] N. Petkov, E. Subramanian, "Motion detection, noise reduction, texture suppression, and contour enhancement by spatiotemporal Gabor filters with surround inhibition," Biol. Cybern., vol. 97, no. 5-6, pp. 423-439, 2007.

[8] Ministry of Land, Infrastructure, Transport and Tourism, Etsumi Mountain Area Sabo Office. http://www.cbr.mlit.go.jp/etsumi/