Change detection using joint intensity histogram

Yasuyo Kita
National Institute of Advanced Industrial Science and Technology (AIST)
Information Technology Research Institute
AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Japan
y.kita@aist.go.jp

Abstract

In the present paper, a method for detecting changes between two images of the same scene taken at different times using their joint intensity histogram is proposed. First, the joint histogram, which is a two-dimensional (2D) histogram of combinatorial intensity levels, $(I_1(\mathbf{x}), I_2(\mathbf{x}))$, is calculated. By checking the characteristics of the ridges of clusters on the joint histogram, clusters that are expected to correspond to background are selected. The combinations of (I_1, I_2) covered by the clusters are determined as insignificant changes. Pixels having a different combinatorial intensity $(I_1(x), I_2(x))$ from these combinations, are extracted as candidates for significant changes. Based on the gradient correlation between the images for each region consisting of these pixels, only regions with significant changes are distinguished. Experiments using real scenes show the practical usefulness of the method.

1. Introduction

Change detection between images of the same scene taken at different times is important in various applications, such as the detection of abnormalities in medical or factory examination, event detection from surveillance images, and change detection from aerial/satellite images [1]. When two images are taken from different view points, the images should be registered first [2]. In the present paper, we concentrate on change detection processes by assuming that the images are taken from the same view point or have already been registered if necessary.

When images are taken consecutively with small time intervals, as with frame-rate video surveillance, the adaptive background subtraction strategy works well when using a series of previous frames[3][4]. However, there are also several situations in which images are taken with a long time interval. In such cases, it is not easy to discriminate "significant" changes, such as the appearance/disappearance of objects, from "insignificant" changes, such as those induced by illumination variation. This difficulty becomes more severe when dealing with outside scenes

In the present paper, we propose a change detection method that estimates significant/insignificant changes based on the joint histogram of two input images. The joint histogram is a two-dimensional (2D) histogram of combinatorial intensity levels, $(I_1(\mathbf{x}), I_2(\mathbf{x}))$, where $I_{1(2)}$ and $\mathbf{x}$ represent the intensity level (0-255) of each image and the positions of pixels on images, respectively. In the case in which a background with insignificant changes is dominant in the images, on the joint histogram, background pixels tend to form large clusters that spread over the (I_1, I_2) plane. By checking the characteristics of the ridges of clusters on the joint histogram, those clusters expected to correspond to background are selected. The combinations of (I_1, I_2) covered by the clusters are determined as insignificant changes. By assuming the remaining combinations to be significant changes, pixels with significant changes are extracted from the images. By checking gradient correlation between the images for each region consisting of sizable connected groups of these pixels, only regions with significant changes are distinguished. In the next section, we describe the proposed method. In Section 3, we present some experimental results to demonstrate the practicality of the proposed method.

2. Methods for detecting changes

2.1. Problems to solve

Before explaining the proposed methods, we need to clarify the problems considered herein. The aim of the study was to detect the appearances/disappearances of objects rather than changes in the intensity of individual objects. We deal with gray-level images in the present paper.

Figures 1(a) and 1(b) provide a typical example in which two images of the same scene do not show simple brightness changes. These images were taken at 7:06 AM and at 11:59 AM, respectively, using a Field Monitoring Server (FMS) camera used as a node of a sensor network[5][6]. Figure 1(c) shows the absolute values of the subtraction of the images considering the ratio of average brightness between the images. That is, $|I_1(\mathbf{x})$-a$I_2(\mathbf{x})|$, where $a = (\sum_{i=1}^{n} I_1(\mathbf{x})/n)/(\sum_{i=1}^{n} I_2(\mathbf{x})/n)$. Here, $I_{1(2)}$, $\mathbf{x}$ and n represent the intensity level (0-255) of each image, the positions of pixels on the images, and the number of pixels in

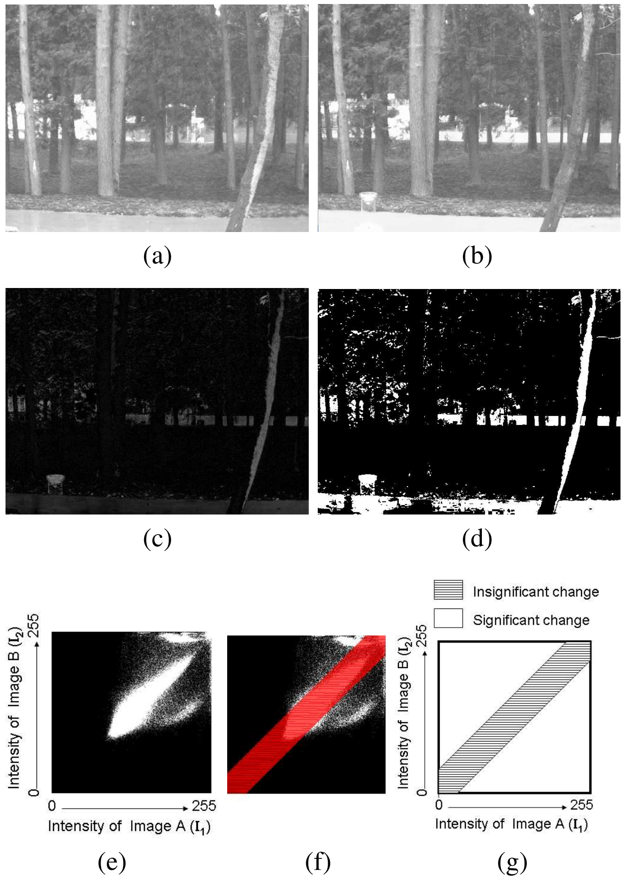


Figure 1. Changed area obtained by fixed threshold

each image, respectively. The white regions in Fig. 1(d) show pixels with changes as obtained by binarization of the subtraction image with the threshold of 30. When the task is the detection of the appearance/disappearance of objects, one FMS camera, which was newly placed on the near side of the road between the image capture times was the only object that should be extracted. However, it is difficult to detect its appearance from the binarization results.

2.2. Clusters on joint histogram

Figure 1(e) shows the joint intensity histogram of the images, in which the frequency is represented by brightness. The binarization after subtraction in the previous subsection can be explained as follows. If $I_1(x)$ is more than $(aI_2(x) - 30)$ and less than $(aI_2(x) + 30)$, then the pixel at $\mathbf{x}$ is assumed to have undergone an insignificant change. This means that the combinations of (I_1, I_2) included in the red region in Fig. 1(f) are assumed to be insignificant changes. In another respect, combinations of (I_1, I_2) are classified into significant/insignificant changes using a 2D table, as shown in Fig. 1(g). Here, we define this table as $Sig(I_1, I_2)$, where $Sig(I_1, I_2) = 1$ and $Sig(I_1, I_2) = 0$ indicate significant and insignificant changes, respectively, and we propose to determine the table more adaptively and flexibly based on the pixel distribution in the joint histogram.

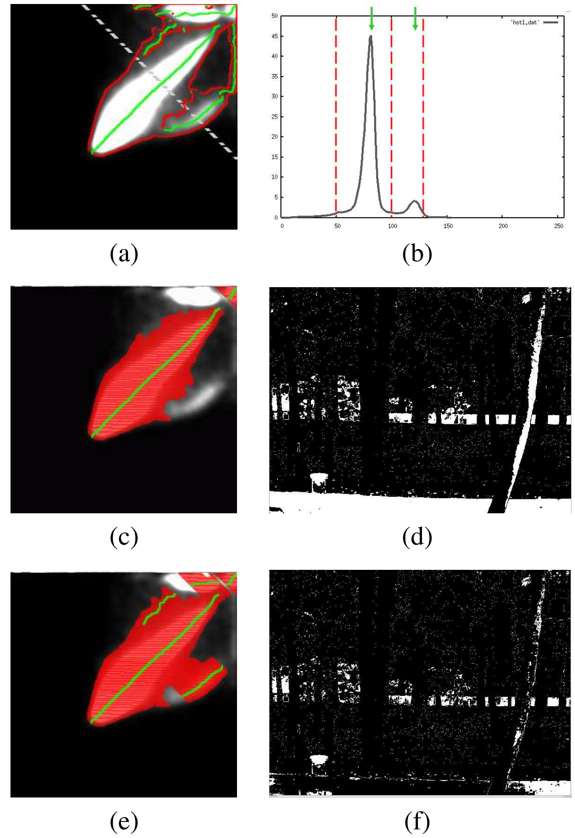


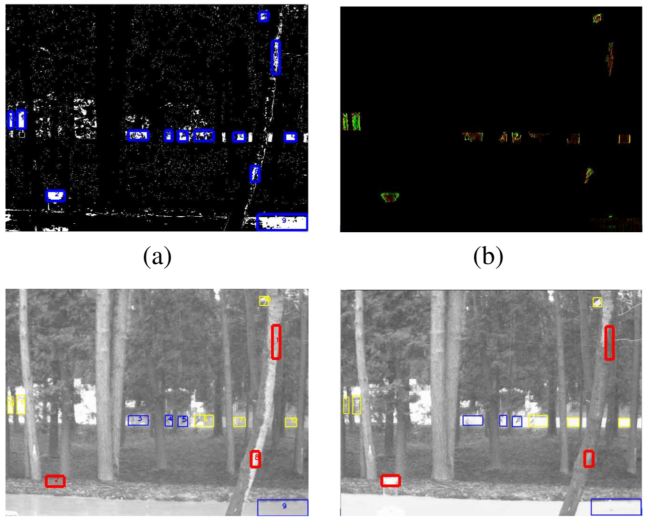
Figure 2. Changed area obtained by Sig(I1,I2)

The joint histogram in Fig. 1(e) has a number of clusters. If there is no change in intensity between two images, then the pixels are plotted on the line $I_2 = I_1$. The global brightness change produces similar shifts of most of pixels. In the case in which pixels affected only by global brightness change are dominant in the images, clusters having long ridge lines are formed. In addition, the projection of the ridge onto the (I_1, I_2) plane has a tangent close to the line $I_2 = I_1$. In contrast, other clusters correspond to pixels having different types of intensity changes that are caused by other factors. In this example, specular reflection of wet roads, changes in the degree of wetness of tree trunks and the appearance of a new object (an FMS camera) are such factors.

Figure 2(a) and 2(b) show the joint histogram smoothed with the 2D Gaussian function of $\sigma = 3.0$ and its cross-section at the $I_2 = -I_1 + 336$, which is illustrated by a dashed line in Fig. 2(a). Peaks and their boundaries are found, as indicated by the arrows and dashed lines in Fig. 2(b). Here, boundaries are determined at the valley between the neighboring peaks or at the position where the value first

Table 1 Gradient similarity (S_e)

region	0	1	2	3	4	5	6	7	8	9	10	11	12
correlation	0.19	0.20	0.05	0.20	0.41	0.31	0.17	0.19	0.07	0.31	0.15	0.02	0.16



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

Figure 3. Discrimination using gradient correlation

becomes less than 1.0. The resultant peaks and its boundaries after the same processing along all cross-sections at $I_2 = -I_1 + b (b = 0 \sim 512)$ are superimposed in Fig. 2(a). The green dots represent peak positions, and the red dots represent boundaries.

If we select only clusters in which the projection of the ridge on the (I_1, I_2) plane is close to $I_2 = I_1$ and assume the combination of (I_1, I_2) included in the clusters to be an insignificant change, then the red regions in Fig. 2(c) indicate the area of $Sig(I_1, I_2) = 0(insignificant)$. Based on the $Sig(I_1, I_2)$ table in which only these combinations are set to 0, pixels with significant changes are extracted, as shown by the white regions in Fig. 2(d). All changes apart from the global intensity change are extracted. The determination of which changes are "significant" depends on the application. However, in many applications, simple changes in the intensities of individual objects, such as roads or tree trunks, are generally required to be ignored. In the following subsection, we show a method by which to select clusters that correspond to such changes.

2.3. Clusters corresponding to insignificant changes

In outside scenes, shading and specular reflection often produce large intensity changes in parts of images. These changes are very difficult to discriminate from significant changes. Under the framework introduced in the previous section, the removal of such intensity changes can be performed by distinguishing clusters that are caused by such factors. When changes in shading (or specular reflection) between two images are not too strong, the pixels affected by the effect maintain similar intensity differences among the pixels that are affected by the same effect. If the number of pixels affected by the same effect is large and these pixels have sufficient variation in intensity level, then the pixels tend to form a cluster that has a long ridge. In addition, the projection of the ridge on the (I_1, I_2) plane has a tangent close to the line $I_2 = I_1$. We hereinafter refer to the slope of the tangent of the projection of a ridge on the (I_1, I_2) plane as the tangent slope of ridge. Using these characteristics, we propose the following algorithm to remove pixels with insignificant changes:

1. Clustering of joint histogram.
   Peaks and their boundaries are detected on the smoothed joint histogram of two input images in the manner described in the previous section. Ridge lines are extracted by connecting adjacent peak points plotted on the (I_1, I_2) plane. The tangent slope of the longest ridge, $D0$, is calculated in order to estimate global brightness change.

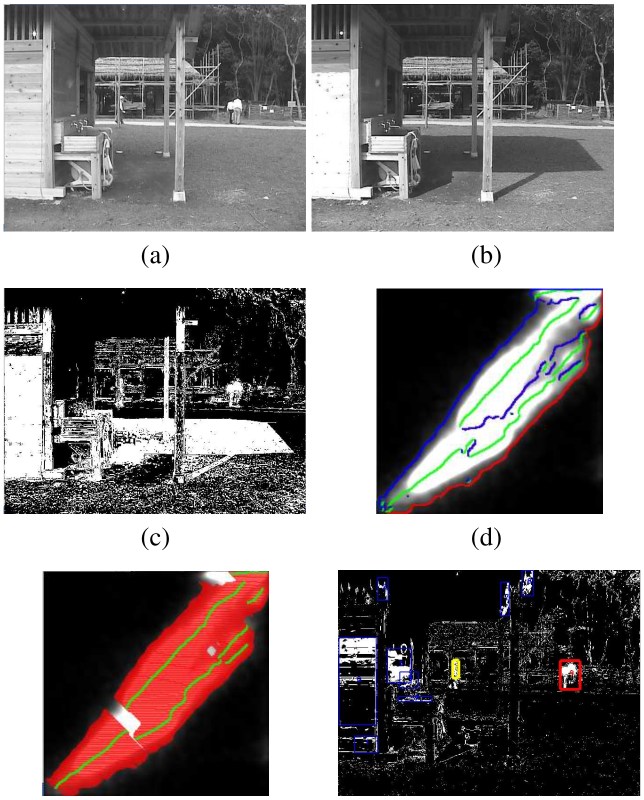
2. Selection of clusters.
   Clusters having ridges longer than $L1$ are selected. From the candidate clusters, only clusters that have a ridge with a tangent slope close to $D0$ (within $\pm D1$ from $D0$) are selected. In order to take the saturation of intensity levels into account, clusters having vertical and horizontal ridges at the top right and bottom left corners are also selected.

3. Extraction of pixels with significant changes.
   The significant/insignificant table, $Sig(I_1, I_2)$, is constructed as follows:

$$\left\{ \begin{array}{ll} Sig(I_1, I_2) = 0 & if\ \ (I_1, I_2)\ \ is\ \ included \\ & in\ \ selected\ \ clusters. \\ Sig(I_1, I_2) = 1 & else \end{array} \right.$$

   Pixels with significant changes are extracted by comparing $(I_1(\mathbf{x}), I_2(\mathbf{x}))$ with the table, $Sig(I_1, I_2)$.

The three parameters σ, $L1$, and $D1$ should be adjusted based on the task and situation for the individual application. However, through experiments using three different scenes, the results were found not to be sensitive to the parameters. In the present paper, with the exception of the



(a) (b)

(c) (d)

(e) (f)

Figure 4. Robustness against large shadow appearance

experiment in Section 3.3, we used the same values of 3.0, 20 (pixels) and 15 (deg) for all of the experiments.

Figure 2(e) shows clusters selected using this algorithm. The red regions show the area of $Sig(I_1, I_2) = 0$. In Fig. 2(f), pixels with significant changes extracted with the table are indicated by white dots. Most of the undesirable changes were removed. As a result, pixels corresponding to the FMS camera, were extracted separately.

2.4. Discrimination using gradient correlation

In the processes in the previous section, insignificant changes that occurred only in small areas are not removed. However, once potential pixels for significant changes are removed from the dominant background, as shown in Fig. 2(f), discrimination of significant changes can be performed through additional filtering processes. We adopt gradient correlation checking for this purpose, and explain the concrete processes using the example of Fig. 3.

Region labeling is applied to images of pixels with significant changes. Regions having more than $N1$ pixels are extracted as candidate regions. In the experiments of the present paper, $N1$ is set to 300 pixels. In the case of Fig. 3(a), the regions surrounded by blue squares are extracted candidate regions. Gradient images of the input images in the x and y directions, $e_{i,k}(i = 1,2; k = x,y)$, are calculated with the Sobel operator. Figure 3(b) shows the gradient image of the x direction in only candidate regions. In the figure, red and green intensities represent the respective gradient strengths of the two input images, resulting in pixels that have similar gradient values in the two images appearing yellow. We define a measure of gradient similarity, S_e, as follows:

$$S_e = min(E_x, E_y)$$
$$E_x = \frac{\sum(e_{1,x}(\mathbf{x}) - \mu_{e_{1,x}})(e_{2,x}(\mathbf{x}) - \mu_{e_{2,x}})}{(max(\sigma_{e_{1,x}}, \sigma_{e_{2,x}}))^2}$$
$$E_y = \frac{\sum(e_{1,y}(\mathbf{x}) - \mu_{e_{1,y}})(e_{2,y}(\mathbf{x}) - \mu_{e_{2,y}})}{(max(\sigma_{e_{1,y}}, \sigma_{e_{2,y}}))^2}$$

where $\mu_{e_{i,k}}$ and $\sigma_{e_{i,k}}$ represent the average and standard distribution of the gradient image in the k direction of image i in a candidate region. The candidate regions are classified into the three classes according to S_e: change with high certainty ($0.0 \leq S_e \leq 0.1$), change with low certainty ($0.1 < S_e \leq 0.3$), and no change ($0.3 < S_e \leq 1.0$)

The partial movement of objects, such as waving trees, often occurs between images, and the detection of such movement is not desirable. When such partial translation occurs, a similar gradient pattern is observed in the vicinity of the corresponding position. Therefore, if we take the maximum of S_e while translating one region over the corresponding region in the other image, such regions should show large S_e (high correlation) and thus can be rejected. The amount of allowable partial translation, $T1$, should be determined by taking into account factors that should be ignored. In the experiments of the present paper, we fixed $T1$ as five pixels.

The values of S_e for the candidate regions in the case of Fig. 3 are shown in Table 1. In Figs. 3(c) and 3(d), the three classes, high certainty, low certainty and no change, are represented using red, yellow and blue squares, respectively. The appearance of an FMS camera was correctly detected, as shown by the red square in the lower left corner. The detection of two other red squares on the tree trunk was undesirable. It is difficult to distinguish these squares based on gradient similarity because the partial wetness of the trunk produces strong new edges.

3. Experiments

We have examined the feasibility of the proposed method using more than 100 images of three scenes taken by FMS cameras. Most of these images were taken using time intervals ranging from two to 30 minutes.

3.1. Robustness against large shadow appearance

Figures 4(a) and 4(b) were captured with a five minute interval between images. These images provide an exam-

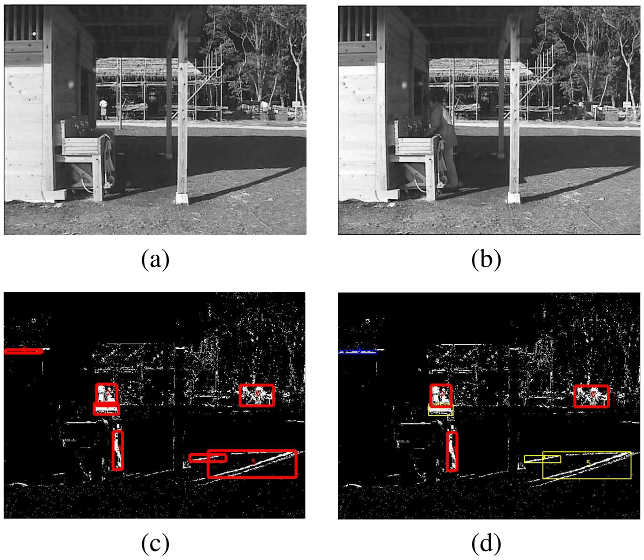


Figure 5. Robustness against partial translation

ple of a rapid change in daylight that causes difficulty with respect to change detection. Figure 4(c) shows the results after binarization of the subtracted image in the same manner as in Fig. 1(d). Figure 4(d) shows the peaks and their boundaries on the joint histogram. Intensity changes due to the appearance of shadows and specular reflections can be observed as separate clusters on the joint histogram. The red regions in Fig. 4(e) show clusters selected by the proposed method. The combinations of (I_i, I_j) covered by the clusters were assumed to be insignificant changes in the table, $Sig(I_i, I_j)$. Based on the table, pixels with significant changes were extracted, as shown in Fig. 4(f). Through the gradient correlation processes, the appearances of people are detected, as indicated by red and yellow squares.

3.2. Robustness against partial translation

Figures 5(a) and 5(b) are successive frames taken at 14:32 PM and 14:42 PM. Between the images, the shadows moved according to the movement of the sun. Since the number of pixels for which the intensity has markedly changed due to this movement is small, the change does not cause a large cluster to form on the joint histogram. Therefore, regions consisting of these pixels were detected as candidates for significant change. If we calculate gradient correlation without the consideration of partial movement, that is, by setting $T1$ 0, all candidates were judged to be significant changes, as indicated by the red squares in Fig. 5(c). However, with an allowability of $T1 = 5$, regions having intensity changes caused by the movement of shadows were discriminated from significant changes. As a result, only the disappearance and appearance of people were correctly detected, as indicated by the red squares in Fig. 5(d).

3.3. Robustness against large daylight change

The four images in the first row of Fig. 6 are scenes of a parking lot taken at one-hour intervals during the evening. Since the scene includes objects of various materials, such as a building, roads, trees and cars, changes in daylight induce various intensity changes. Therefore, these present very difficult cases for change detection. The variety of intensity changes can be seen as the scattering of pixels on the joint histograms (second row of Fig. 6). The third row shows the change detection results obtained by the proposed method. Although a few regions were also detected undesirably, the proposed method managed to extract most of appearances/disappearances of cars, except in the final case. The reason for this is that, in this case, many background pixels appear as small clusters on the joint histogram and their clusters were not selected. As a result, pixels corresponding to newly appearing cars were not extracted separately. The forth row in Fig. 6 shows the results after changing parameters to select smaller background clusters, specifically, $L1 = 10$ and $D1 = 30$. With these new parameters, the detection of the appearance of a car was possible in the final case. However, taking smaller clusters increases the risk of missing significant changes.

4. Conclusion

We proposed a method for detecting changes between two images of the same scene taken at different times using the joint intensity histogram of the images. Although Bromiley[7] also used joint histogram (scattergrams) to transform one image as preprocessing of image subtraction in medical applications, the approach proposed herein uses more information from the joint histogram and directly classifies combinations of (I_1, I_2) into significant/insignificant changes. This flexibility in determining significant/insignificant changes and the adaptability to each pair of input images allows the proposed method to achieve good performance in difficult cases of change detection from outside scenes. Although the proposed method barely extracts all pixels of regions with significant change, especially in cases of large changes in daylight, we believe that post-processing can be used to compensate for this effect by using the pixels extracted by this method as seeds.

Note that on the joint histogram, the image coordinates of pixels are ignored. If a significant change happens to produce the same intensity change as that of the background, the proposed method can not distinguish it and so removes all pixels showing the intensity change. From experiments using approximately 300 pairs of images, such coincidences did not occur as often as expected. Only one disappearance of a person between images was missed, even though the change was fairly clear. We are investigating a method by

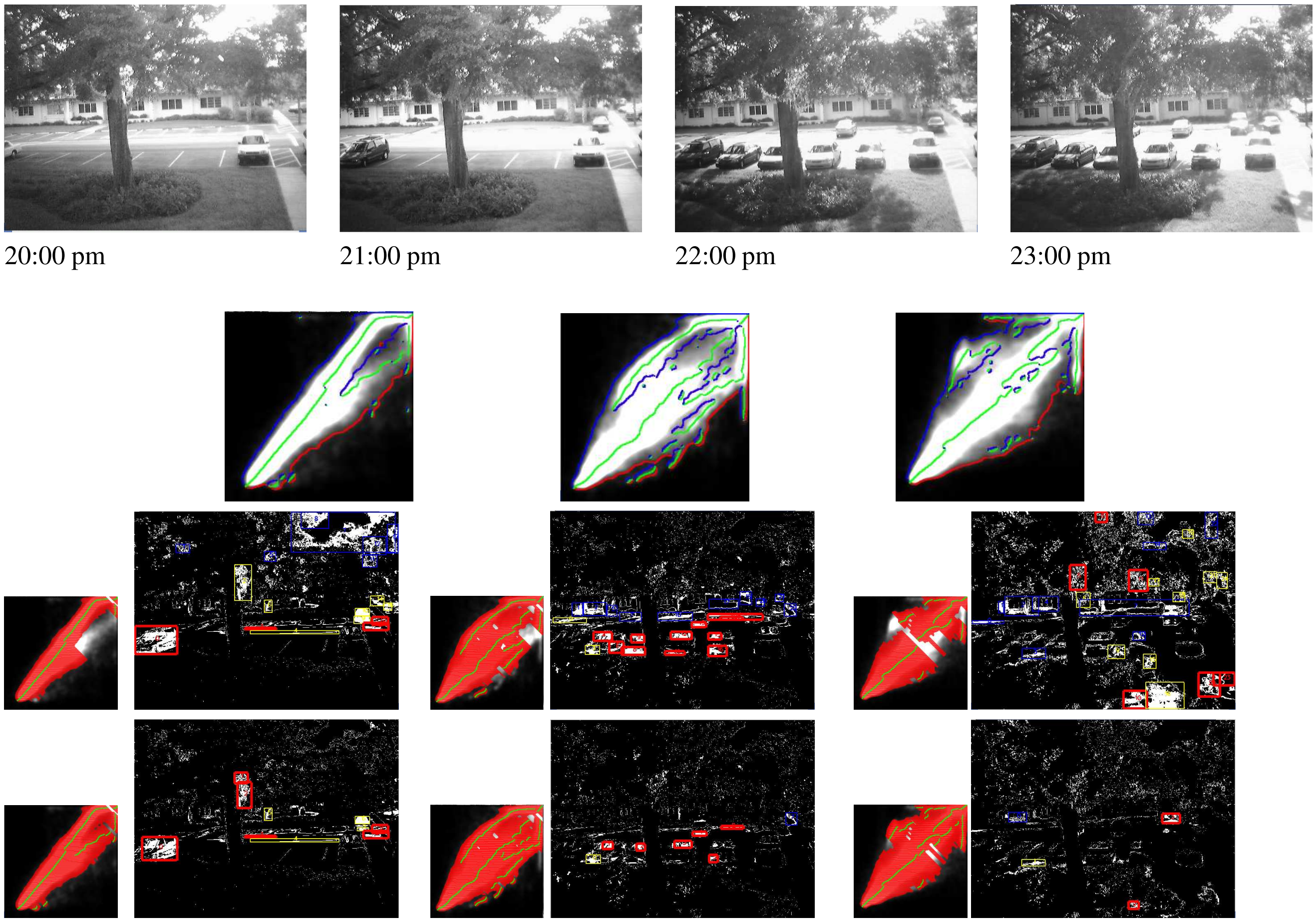


Figure 6. Robustness against large daylight change

which to combine geometric information with the joint histogram in order to overcome this limitation.

Although, in the present paper, we have described experiments that were conducted using only surveillance camera images, we believe that the proposed method is applicable to other applications, such as changes in the detection of aerial/satellite images by using task-oriented rules to select clusters that correspond to insignificant changes.

Acknowledgments

We are thankful to Dr. S. Ninomiya and Dr. M. Hirafuji of National Agricultural Research Center for data provision and discussions, and also to Dr. K. Sakaue, Dr. T. Nagami, Dr. T. Masuda, and Dr. N. Kita for their support to this research.

References

[1] R. J. Radke, S. Andra, O. Al-Kofahi and B. Roysam: "Image change detection algorithms: a systematic survey", *IEEE Trans. on Image Processing*, Vol. 14, No.3, pp. 294–307, 2005.

[2] B. Zitová and J. Flusser: "Image registration methods: a survey", *Image and Vison Computing*, Vol. 21, No. 11, pp. 977–1000, 2003.

[3] K. Toyama, J. Krumm, B.Brumit, and B. Meyers: "Wallflower: principles and practice of background maintenance", In *Proc. of International Conference on Computer Vision*, pp. 255–261, 1999.

[4] M. Pic, L. Berthouze, and T. KuritaB: "Adaptive background estimation: computing a pixel-wise learning rate from local confidence and global correlation values", *IEICE Trans. on Information and System*, Vol. E87-D, No. 1, pp. 50–57, 2004.

[5] http://model.job.affrc.go.jp/FieldServer/FieldServerEn/default.htm:

[6] T. Fukatsu and M. Hirafuji: "Field Monitoring Using Sensor-Nodes with a Web Server", *Journal of Robotics and Mechatronics*, Vol. 17, No.2, pp. 164–172, 2005.

[7] P.A. Bromiley, N.A. Thacker, and P. Courtney: "Non-parametric Image Subtraction using Grey Level Scattergrams", *Image and Vison Computing*, Vol. 20, pp. 609–617, 2002.