2013 IEEE Conference on Computer Vision and Pattern Recognition

# Background Modeling based on Bidirectional Analysis

Atsushi Shimada, Hajime Nagahara, Rin-ichiro Taniguchi
Graduate School of Information Science and Electrical Engineering, Kyushu University
744, Motooka, Nishi-ku, Fukuoka, 819–0395 Japan

## Abstract

*Background modeling and subtraction is an essential task in video surveillance applications. Most traditional studies use information observed in past frames to create and update a background model. To adapt to background changes, the background model has been enhanced by introducing various forms of information including spatial consistency and temporal tendency. In this paper, we propose a new framework that leverages information from a future period. Our proposed approach realizes a low-cost and highly accurate background model. The proposed framework is called bidirectional background modeling, and performs background subtraction based on bidirectional analysis; i.e., analysis from past to present and analysis from future to present. Although a result will be output with some delay because information is taken from a future period, our proposed approach improves the accuracy by about 30% if only a 33-millisecond of delay is acceptable. Furthermore, the memory cost can be reduced by about 65% relative to typical background modeling.*

## 1. Introduction

Background modeling and subtraction is an essential task in video surveillance applications, as it provides foreground segmentation with no prior information about the foreground. Pixel-level background modeling is a typical approach in which a Gaussian mixture model (GMM) or kernel density estimation is often used to represent the frequency of pixel values in an observed image sequence[14, 4]. Region-level background modeling is also often studied. Instead of using pixel values, features extracted from relationships between a central pixel and its surrounding pixels are used. Compared with pixel-level modeling, region-level modeling gives richer image features and is robust in the case of illumination changes[8, 17]. Other effective solutions to enhance the performance of background subtraction are the use of temporal information[13, 18] and hybrid modeling[16, 15].

The above approaches are common in that they use information taken from past image frames, as drawn on the left side of Figure 1. In contrast, our approach focuses on information that will be observed in future image frames. Generally, future information is not often used in time-series analysis that requires real-time processing since there is a delay in the availability of a result. Practically, however, there is an "acceptable delay" depending on the application. For instance, visual surveillance applications aim to observe the real world and detect particular events as quickly as possible. In such a case, although it is desirable to output a result with real-time processing, a slight delay in output will not introduce difficulties.

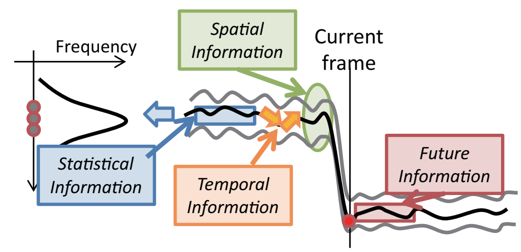



Figure 1. Statistical, spatial and temporal information used in traditional approaches, and future information used in the proposed approach. The solid black line and the solid gray lines represent the pixel of interest and its peripheral pixels respectively.

Our approach defines an acceptable delay as 33 milliseconds (the duration of just one video frame). The background model is improved in terms of its ability to handle background changes and accurately subtract the background compared with a typical approach that does not use future information. Moreover, our approach obtains the background model using the same amount of memory as used in the typical approach even though it uses additional information obtained from future image frames.

## 2. Bidirectional Background Modeling

### 2.1. Concept

The proposed method is based on two concepts.

**Acceptable Delay** The background model is allowed to output a result $N$ frames after the current frame. In

1063-6919/13 $26.00 © 2013 IEEE
DOI 10.1109/CVPR.2013.258



IEEE computer society

other words, information observed in the period extending to $N$ frames after the current frame is used to determine the background subtraction. The proposed method improves the accuracy of background subtraction at the expense of a delay in the output. However, the proposed method requires a delay of just one frame, which can be ignored in most visual surveillance applications.

**Backward Analysis** Generally, time-series analysis is performed in the forward direction; i.e., from the past to present. In contrast, the proposed method includes backward analysis using $N$ future frames, in addition to forward analysis. The backward analysis is performed from the future to present. Figure 2 shows a typical example of the advantage realized by including backward analysis. The left side of Figure 2 illustrates situations when a pixel value suddenly changes. The upper case is the result of an illumination change, and the lower case is the result of an object moving. These changes are the same from the viewpoint of the change in the pixel value, and background modeling based on forward analysis cannot distinguish the reason for the changes at the time of the current frame. In contrast, backward analysis is able to investigate the change using the pixel values observed in the future period (the right side of Figure 2). If the change is due to a moving object, both forward and backward analyses will observe a change in the pixel value. In contrast, if the change is due to an illumination change, only one of the analyses will observe a change in the pixel value.

Here, we recall the characteristics of the proposed method. The proposed bidirectional background model uses pixel values observed in the future period to improve the accuracy of background subtraction at the expense of some delay in the output. However, we do not require a long delay. We are able to acquire a result for background subtraction with a reasonable delay. Therefore, the proposed method is a practical framework for surveillance applications.

## 2.2. Formulation

This section presents a formulation strategy to clarify the proposed method. For simplicity, the explanation focuses on a certain pixel. Let $X^t$ be a pixel value in frame $t$ that represents an observed sequence $\{X^0, \dots, X^{t-1}\}$. In the case of forward analysis, a background model $M^{t-1}$ is estimated from this sequence, and whether an observed pixel value $X^t$ is part of the background is determined by $P(X^t|M^{t-1})$. (In fact, a GMM-based background model is often used for the calculation of background probability.) Meanwhile, backward analysis provides a background model $M^{t+1}$ using $\{X^{t+N}, \dots, X^{t+1}\}$ to calculate the

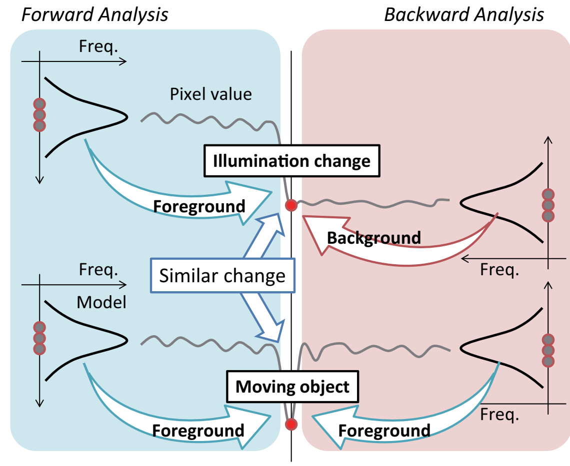



Figure 2. The advantage of including backward analysis. The upper part shows an example of a change in pixel value due to illumination change. The lower part is the case of a moving object.

background probability $P(X^t|M^{t+1})$. In the remainder of this paper, we refer to the background model $M^{t+1}$ as the "backward background model".

The proposed bidirectional background modeling can be said to calculate the background probability of $X^t$ as $P(X^t|M^{t-1}, M^{t+1})$, where $M^{t-1}$ and $M^{t+1}$ are acquired by forward analysis and backward analysis respectively. The probability is given by

$$P(X^t|M^{t-1}, M^{t+1}) = (1-\alpha)P(X^t|M^{t-1}) + \alpha P(X^t|M^{t+1}), \quad (1)$$

where $\alpha, (0 \le \alpha \le 1)$ is a parameter that adjusts the contribution of the backward background model. Note that if we set $\alpha$ to zero, the model is a typical background model based on forward analysis alone.

## 2.3. Backward Background Model

Ideally, the acceptable delay $N$ should be a large value to acquire a good backward background model. Meanwhile, $N$ should be as small as possible for practical use. To solve this trade-off, a new concept of "piecewise time-reversal symmetry" is introduced, where we can set $N$ to a small value yet realize a reasonable backward background model.

Piecewise time-reversal symmetry is an assumption that background change has a symmetric property in a short period if the order of observation from past to present is inversed to observation from future to present. For instance, phenomena of "a pixel getting darker" and "a pixel getting brighter" are symmetric. A phenomenon of "a pixel getting brighter then getting darker repeatedly" also has a symmetric property if we consider a short time period of the repetition. Figure 3 shows an example of piecewise time-reversal symmetry. If we inverse the piecewise change within period B, the inversed sequence appears similar to the sequence

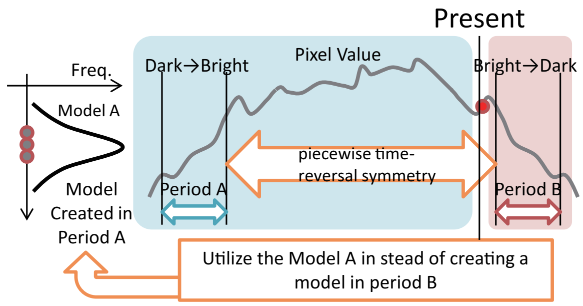



Figure 3. An illustration of piecewise time-reversal symmetry

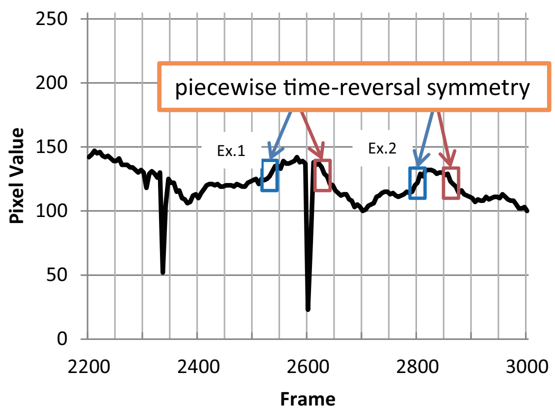



Figure 4. Examples of piecewise time-reversal symmetry. The horizontal axis is the frame number and the vertical axis is the pixel value.

within period A. If we can assume this kind of time-reversal symmetry, an observed sequence of pixel values in the past period might include a time-reversal pattern that will be observed in the future period. A background model created for the piecewise past period could then be used on behalf of a background model that is estimated by pixel values in the future period. In other words, we do not have to explicitly create a backward background model since we can substitute an "inversed forward" background model for a backward background model.

The implementation is described in more detail in section 3. Here, we illustrate the relevancy of piecewise time-reversal symmetry with an actual observation of pixel values in Figure 4. The sequence of pixel values is taken from an outdoor scene. The piecewise change within the blue rectangle is time-reversal symmetric for the one within the red rectangle. These types of symmetric changes were often observed in other scenes used in our experiments.

## 3. Implementation

This section explains the detailed implementation strategy to apply the proposed bidirectional background modeling using a GMM-based statistical background model.

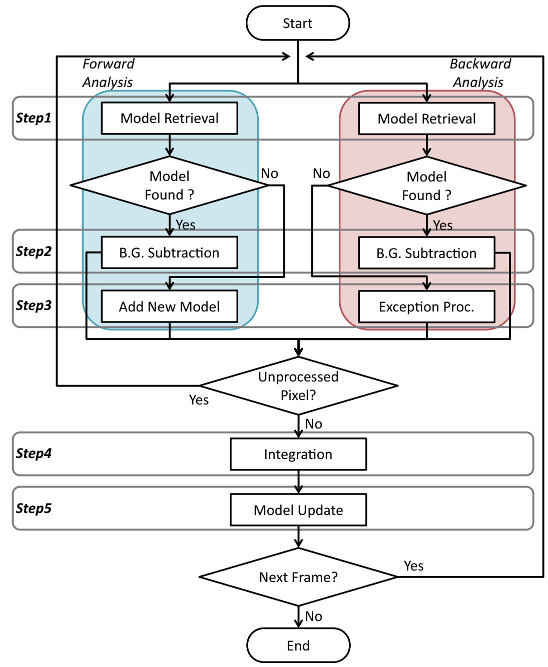



Figure 5. Flow chart of the proposed method

### 3.1. Process Flow

Figure 5 depicts the process flow of the proposed method. To explain the process, we focus on a certain pixel $i$ in an image. The same process is performed for each pixel. For simplicity, the notation $i$ is ignored in the following explanation. In addition, the notation with regard to time $t$ is also ignored since the process from Step 1 through Step 5 is performed for the same frame.

**Step 1: Background model retrieval** A background model $M^q$ or $M^r$ that satisfies a search query $q$ for forward analysis or $r$ for backward analysis is retrieved from the background database (which is described in detail in section 3.2). Each query $q$ or $r$ is constructed from a pixel feature in the past period or future period respectively. The actual design of the query is described in section 3.2.

**Step 2: Background subtraction** If a background model is retrieved in Step 1 (i.e., $M^q$ and/or $M^r$ is found in the database), a label candidate (foreground or background) for the current pixel value $X$ is estimated as $P(X|M^q)$ and/or $P(X|M^r)$. According to GMM-based background modeling, if the pixel value $X$ is within a predefined standard deviation $s$ of the distribution, the background label is tentatively given to the pixel. Otherwise, it is labeled as foreground.

**Step 3: Adding a new example and exception processing** If a background model is not retrieved, the process is a little different between the foreground analysis and backward analysis. In the case of forward analysis in Step 1, a new GMM-based background model is added to the database with initial mean value $X$ and predefined variance and weight. In this case, the foreground label is tentatively given to the pixel since there is no example that guarantees the pixel to be background. In the case of backward analysis, we only give a tentative label of foreground to the pixel, and do not add any background model to the database.

**Step 4: Integration** After processing the above three steps for all pixels, final labels are assigned using tentative labels given in Steps 2 and 3. Section 3.3 gives a detailed explanation.

**Step 5: Update of background models** The parameters of the background models are updated. The update process is applied for retrieved backgrounds only (i.e., background models used in Step 2). Note that when a background model is used by more than one pixel, one of the pixels is randomly selected for the update.

### 3.2. Case-based Background Model Retrieval

This section gives a detailed explanation of Step 1 in Figure 5. "Case-based background model retrieval" is a framework with which to realize "case-by-case model sharing". Unlike the clustering-based approach or traditional pixel-based approaches, the same background model is not continuously used for an individual pixel. Instead, according to a condition of the observed pixel value (e.g., the location of the pixel or the trend of the value), an appropriate background model is selected from the database for an individual pixel frame by frame, meaning that a given background model is not always selected for the same pixel. Moreover, a background model is sometimes shared by several pixels.

The important point is that we do not create separate background databases for forward analysis and backward analysis. Forward and backward analyses share the same background database through the use of piecewise time-reversal symmetry. In practice, the query to retrieve a background model is set as follows.

$$q = (X^{t-1}, X^t, u, v) \tag{2}$$

$$r = (X^{t+1}, X^t, u, v) \tag{3}$$

where $X^t$ is a pixel value in frame $t$, and $(u, v)$ are the two-dimensional coordinates of the pixel.

In forward analysis, a background model $M^q$ is retrieved where a similar pixel change was observed around $(u, v)$ in the past period. The retrieved $M^q$ is regarded as $M^{t-1}$ in Eq. (1). On the other hand, the query of the backward analysis changes the time ordering of pixel values. $X^{t+1}$ is followed by $X^t$ (from future to present) in the case of query $r$. Therefore, a background model that corresponds to piecewise time-reversal symmetric change will be retrieved as $M^r$ from the database. $M^r$ is also regarded as $M^{t+1}$ in Eq. (1).

### 3.3. Foreground/Background Label Assignment

Each pixel has two tentative labels from the forward analysis and backward analysis. The final label is assigned by integrating the tentative results with consideration of consistency with adjacent pixels.

An energy function is defined according to a Markov random field and each pixel is given a proper label (foreground or background) by minimizing the energy function. The energy function is defined as

$$E(\boldsymbol{L}|\boldsymbol{X}) = \lambda \sum_{i \in \mathcal{V}} G(l_i|X_i) + \sum_{(i,j) \in \mathcal{E}} H(l_i, l_j|X_i, X_j), \tag{4}$$

where $L = (l_1, \ldots, l_M)$ is a binary label vector, and $M$ is the number of pixels. $\mathcal{V} and \mathcal{E}$ represent a set of all pixels and a set of four adjacent pixels respectively. $G(l_i|X_i)$ and $H(l_i, l_j|X_i, X_j)$ represent the penalty term and smoothing term respectively and are calculated as

$$G(l_i|X_i) = (1-\alpha)P(X_i|M^q) + \alpha P(X_i|M^r) \tag{5}$$

$$H(l_i, l_j|X_i, X_j) = \frac{1}{\ln(|X_i - X_j| + 1 + \epsilon)}. \tag{6}$$

The energy is minimized using a graph-cut algorithm[2].

### 3.4. Gaussian Mixture Background Model

This section briefly explains Gaussian mixture background models. The probability of observing the current pixel value $X^t$ is

$$P(X^t) = \sum_{k=1}^{K} w_k^t \eta(X^t|\mu_k^t, \boldsymbol{\Sigma}_k^t), \tag{7}$$

where $K$ is the number of distributions. The variables $w_k^t$, $\mu_k^t$ and $\boldsymbol{\Sigma}_k^t$ are an estimate of the weight, mean value and covariance matrix of the $k$-th Gaussian in the mixture for frame $t$, respectively, while $\eta$ is the Gaussian probability density function. Each parameter is updated to adapt to an observed pixel value frame by frame. According to the change in pixel value, the number of distributions changes dynamically. For further details, refer to [12].

## 4. Experimental Results

### 4.1. Preparation

The evaluation items in our experiments are the accuracy of background subtraction, memory cost and computational time. The accuracy was evaluated using the precision ratio,

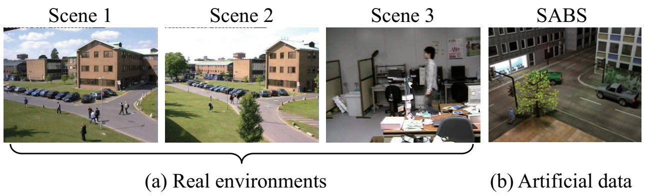



Figure 6. Evaluated scenes

recall ratio and F-measure given by

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP + FN}}, \tag{8}$$

$$\text{F} = 2/\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right). \tag{9}$$

The F-measure quantifies the balance between precision and recall, and a larger value reflects a better result. TP, FP and FN denote the number of pixels detected correctly, detected incorrectly, and undetected incorrectly, respectively.

Open datasets were used for evaluation; three real scenes and six artificial scenes (Stuttgart Artificial Background Subtraction Dataset: SABS[3]). The real scenes (see Figure 6(a)[1]) include two outdoor scenes and one indoor scene where illumination changes often occur. The artificial datasets (see Figure 6(b)) separately include the following background changes.

**Basic** The scene includes basic background changes (e.g., waving trees)

**Dynamic Background** Some parts of the scenery may contain movement, but should be regarded as background, according to their relevance. Such movement can be periodic or irregular (e.g., traffic lights and waving trees).

**Bootstrap** If initialization data free from foreground objects are not available, the background model is initialized using a bootstrapping strategy.

**Darkening** It is desirable that the background model adapts to gradual changes in the appearance of the environment. In outdoor settings, for example, the light intensity typically varies during the day.

**Light Switch** Sudden one-off changes are not covered by the background model. They occur, for example, with a sudden switch of light, and they strongly affect the appearance of the background and result in false positive detections.

**Noisy Night** A video signal is generally superimposed with noise. Background subtraction approaches for video surveillance have to cope with such degraded signals affected by different types of noise, such as sensor noise and compression artifacts.

[1] The ground truth is available at http://limu.ait.kyushu-u.ac.jp/dataset/

Table 1. Background subtraction accuracy for outdoor scenes

| | | Scene 1 | Scene 2 | Scene 3 |
|---|---|---|---|---|
| GMM[14] | Precision | 0.26 | 0.23 | 0.07 |
| | Recall | **0.74** | 0.68 | **0.90** |
| | F-Measure | 0.39 | 0.34 | 0.13 |
| Case-based | Precision | 0.60 | 0.43 | 0.16 |
| | Recall | 0.69 | 0.69 | 0.82 |
| | F-Measure | 0.64 | 0.52 | 0.27 |
| Proposed | Precision | **0.72** | **0.58** | **0.49** |
| | Recall | 0.71 | **0.70** | 0.51 |
| | F-Measure | **0.72** | **0.64** | **0.50** |

With regards to parameter settings, we set the contribution parameter $\alpha$ to $0.5$, which was determined from preliminary experiments. Other parameter settings except for the standard deviation $s$ were taken from the literature[12]. The standard deviation $s$ was changed within the range of $[0.5, 250]$ to draw precision and recall charts.

### 4.2. Accuracy of Background Subtraction in Real Scenes

Three methods, namely the GMM-based method[14], case-based method (i.e., the parameter $\alpha = 0$) and the proposed method, were compared in terms of background subtraction accuracy. Figure 7 shows foreground masks for each scene and Table 1 gives the evaluation results for the maximal F-measure[2]. A value given in bold type is the best score among the three methods.

The result of the GMM-based method[14] is a typical baseline with much lower precision and higher recall because of the method's low flexibility to background changes. The case-based method (which did not employ backward analysis) provides better results than the GMM-based method. The proposed method further improved accuracy, especially for the indoor scene (Scene 3). The indoor scene includes sudden illumination changes caused by the switching of a light on/off. In such scenes, the advantage of backward analysis in the proposed method is demonstrated as illustrated in Figure 2.

The GMM was likely to detect false positive pixels, resulting in low precision. The proposed method could reduce such false positives by bidirectional analysis, i.e. using two GMMs (forward and backward). Therefore, the backward analysis hypothesis contributed to gain the accuracy.

### 4.3. Memory and Computational Costs

The memory usage and computational time were also investigated. With regard to memory usage, the amount of memory used by the program code was also monitored. The computational time was recorded frame by frame, and the average time was evaluated. Table 1 gives the results. The

[2] MRF smoothing was used for GMM[14] as well as other methods.

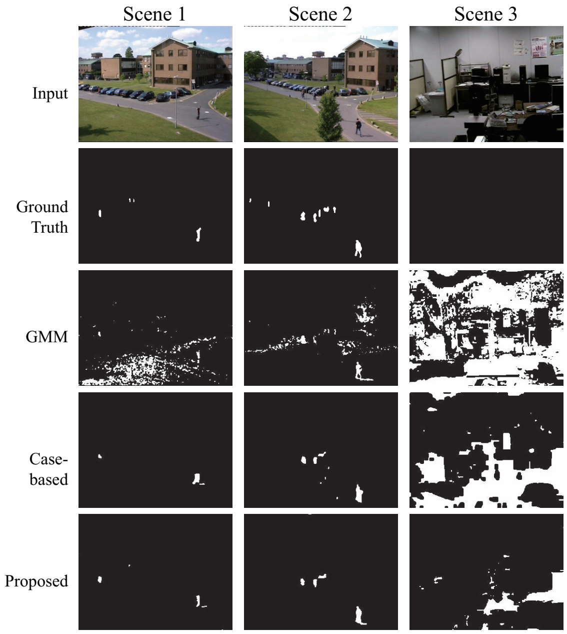



Figure 7. Foreground masks

Table 2. Memory cost and computational time

| | Memory | Time |
|---|---|---|
| GMM[14] | 1 | 1 |
| Case-based | **0.30** | **0.32** |
| Proposed | 0.33 | 0.35 |

values in the table indicate the ratio to the baseline; i.e., the GMM-based method.

The case-based method had both lower memory cost and lower computational time than the GMM-based method. These good results were a consequence of the "case-by-case model sharing" strategy employed, where several pixels shared the same model. This reduced the total number of background models. As a result, the computational time was also reduced. The ratios of the proposed method are almost the same as those of the case-based method because the same background database was used even though the proposed method employed analyses in two directions.

## 4.4. Adequacy of Piecewise Time-reversal Symmetry

The proposed method assumed piecewise time-reversal symmetry. We investigated the validity of the assumption by comparing the proposed method with full-backward analysis. The backward background model was completely estimated using all the frames from the end of the image sequence to the initial frame. The same integration strategy (i.e., graph-cut-based label assignment as mentioned in section 3.3) was then applied to estimate the foreground/background label. Therefore, the full-backward analysis was performed in off-line processing. Almost the same results were acquired by the full-backward analysis as

Table 3. Background subtraction accuracy when using all the frames for backward analysis

| | | Scene 1 | Scene 2 | Scene 3 |
|---|---|---|---|---|
| full- | Precision | 0.71 | 0.67 | 0.42 |
| backward | Recall | 0.70 | 0.64 | 0.57 |
| analysis | F-Measure | 0.70 | 0.65 | 0.48 |

shown in Table 3 compared with the results of the proposed method in Table 1. Therefore, the assumption is basically justified.

We furthermore conducted an additional experiment with short-term future sequences. The result was almost the same with the full backward analysis. We suppose that the background model update is strongly affected by the successive frames from the current frame even using all of the future sequences.

## 4.5. Evaluation with an SABS dataset

This section reports evaluation results obtained using an SABS dataset. In the literature [3], nine approaches including recent methods and more conventional methods have been evaluated for several scenes. Figure 8 shows the recall–precision curves when changing the parameter $s$ as mentioned in section 4.1. Additionally, Table 4 gives the results for the maximal F-measure for each scene.

Overall, the accuracy of the proposed method was higher than that of other methods. According to the maximal F-measure in Table 4, the proposed method achieved the top score for three out of six scenes. The average score for all scenes was also higher than the average scores for the other methods. Therefore, we can argue that the proposed method has broad utility.

Considering each scene in turn, all methods achieved high scores for the scenes "Basic" and "Dynamic Background" since these scenes did not include severe background changes. Although the proposed method did not achieve the highest scores for these scenes, the results compared favorably with the results obtained with other methods. The proposed method did not provide high accuracy for the "Darkening" scene, achieving a result less than $60\%$. The reason for this is that the scene included the background getting darker only. The inverse change of the background getting brighter was not included, and therefore, the assumption of piecewise time-reversal symmetry did not work well. This is a limitation of the proposed method; however, considering the practical use, background subtraction is usually applied to scenes that not only become darker but also become brighter. Indeed, the proposed method performed well for the real scene (Scene 1, captured outdoors), which included the background becoming darker.

The proposed method outperformed other methods for the scenes of "Bootstrapping", "Light Switch" and "Noisy Night". With regard to "Bootstrapping", we suppose that the model-sharing strategy rather than the bidirectional analysis contributed to the high accuracy. As the training frames are not available in “Bootstrapping”, typical approaches suffer from background initialization and updating of the model. Meanwhile, the case-based sharing strategy used in the proposed method can tackle such an initialization problem by creating a new background model immediately.

In the cases of “Light Switch” and “Noisy Night”, the proposed backward analysis contributed considerably to an improvement in accuracy. Especially in the scene of “Light Switch”, as well as Scene 3 captured indoors, the effectiveness of the proposed method was confirmed.

Finally, the literature does not discuss the implementation cost of the nine compared methods. However, the implementation of the compared methods will have memory costs much greater than the cost of implementing the proposed method. Therefore, the proposed method provides a good balance of high accuracy and low cost.

## 5. Discussion

The experiments described above clarified the advantages of the proposed method. Firstly, the proposed method achieved better results than most other methods including state-of-the-art methods in terms of the accuracy of background subtraction in various scenes. The experimental results supported the effectiveness of bidirectional analysis.

Secondly, the proposed method can be implemented with low-cost memory usage. There are two factors that contribute to the reduction in memory cost. One is the case-by-case model sharing strategy, which allows pixels to share a background model according to the pixel property. The other is the idea that piecewise time-reversal symmetry allows forward analysis and backward analysis to share the same background database. As a result, the proposed method has reduced memory cost suitable for practical application.

Finally, the accepted delay for our concept was set to just one frame in our implementation[3]. The delay is 33 milliseconds in the case of a video sequence. The 33-millisecond delay can be ignored in most visual surveillance applications, but it improves the accuracy of background subtraction remarkably.

## 6. Conclusion

This paper discussed background modeling based on bidirectional analysis. The introduction of backward analysis and its combined use with forward analysis provide a good solution to improve background subtraction accuracy. The proposed framework can be realized with low-memory cost and short computational time, which are suitable for practical use. Experimental results using real scenes and artificial scenes demonstrated these advantages of the proposed method.

The proposed method still has some limitation that the backward analysis does not always work well in some scenes where the pixel values constantly increase/decrease, where the occlusion lasts for a long time, including the situations of a human/car stops, near-field object detection. Meanwhile, it is also true that the effectiveness of the proposed method could be confirmed in far-field sensing datasets evaluated in this paper.

In future work, further scenes need to be used in evaluating the proposed method, and application of the bidirectional background modeling framework to other background models will be studied. We believe that the case-based background modeling framework has great potential.

[3] See the query $r$ in Eq. (2), which uses only $X^{t+1}$ as future information.

## References


[1] O. Barnich and M. V. Droogenbroeck. Vibe: A powerful random technique to estimate the background in video sequences. In *In IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pages 945–948, 2009.

[2] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1124–1137, 2004.

[3] S. Brutzer, B. Hoferlin, and G. Heidemann. Evaluation of background subtraction techniques for video surveillance. In *Computer Vision and Pattern Recognition (CVPR)*, pages 1937–1944. IEEE, 2011.

[4] A. Elgammal, R. Duraiswami, D. Harwood, and L. Davis. Background and Foreground Modeling using Non-parametric Kernel Density Estimation for Visual Surveillance. *Proceedings of the IEEE*, 90:1151–1163, 2002.

[5] K. Kim, T. Chalidabhongse, D. Harwood, and L. Davis. Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11(3):172–185, 2005.

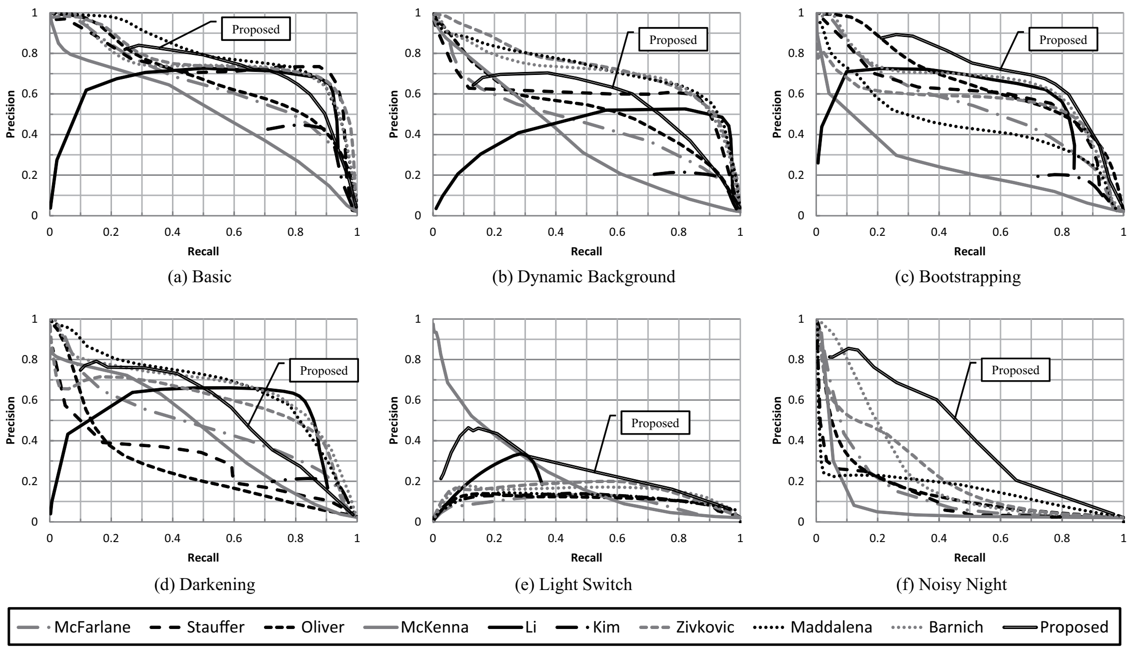
[6] L. Li, W. Huang, I. Gu, and Q. Tian. Foreground object detection from videos containing complex background. In *In Int. Conf. on Multimedia*, pages 2–10, 2003.

[7] L. Maddalena and A. Petrosino. A self-organizing approach to background subtraction for visual surveillance applications. *IEEE Transactions on Image Processing*, 17(7):1168–1177, 2008.

[8] H. Marko and P. Matti. A Texture-Based Method for Modeling the Background and Detecting Moving Objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):657–662, 2006.

[9] N. McFarlane and C. Schofield. Segmentation and tracking of piglets in images. *Machine Vision and Applications*, 8(3):187–193, 1995.

[10] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler. Tracking groups of people. *Computer Vision and Image Understanding*, 80(1):42–56, 2000.




(a) Basic (b) Dynamic Background (c) Bootstrapping

(d) Darkening (e) Light Switch (f) Noisy Night

Figure 8. Precision and recall for the SABS dataset

Table 4. Maximal F-measures for the SABS dataset

| Method | Basic | Dynamic Background | Bootstrap | Darkening | Light Switch | Noisy Night | Scene Average |
|---|---|---|---|---|---|---|---|
| McFarlane[9] | 0.614 | 0.482 | 0.541 | 0.496 | 0.211 | 0.203 | 0.425 |
| Stauffer[14] | **0.8** | 0.704 | 0.642 | 0.404 | 0.217 | 0.194 | 0.494 |
| Oliver[11] | 0.635 | 0.552 | – | 0.3 | 0.198 | 0.213 | 0.380 |
| McKenna[10] | 0.522 | 0.415 | 0.301 | 0.484 | 0.306 | 0.098 | 0.354 |
| Li[6] | 0.766 | 0.641 | 0.678 | **0.704** | 0.316 | 0.047 | 0.525 |
| Kim[5] | 0.582 | 0.341 | 0.318 | 0.342 | – | – | 0.396 |
| Zivkovic[19] | 0.768 | 0.704 | 0.632 | 0.62 | 0.3 | 0.321 | 0.558 |
| Maddalena[7] | 0.766 | **0.715** | 0.495 | 0.663 | 0.213 | 0.263 | 0.519 |
| Barnich[1] | 0.761 | 0.711 | 0.685 | 0.678 | 0.268 | 0.271 | 0.562 |
| Proposed | 0.723 | 0.623 | **0.708** | 0.577 | **0.335** | **0.475** | **0.574** |


[11] N. Oliver, B. Rosario, and A. Pentland. A bayesian computer vision system for modeling human interactions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):831, 2000.

[12] A. Shimada, D. Arita, and R. ichiro Taniguchi. Dynamic Control of Adaptive Mixture-of-Gaussians Background Model. *CD-ROM Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2006.

[13] A. Shimada and R. Taniguchi. Hybrid Background Model using Spatial-Temporal LBP. In *IEEE International Conference on Advanced Video and Signal based Surveillance 2009*, 09 2009.

[14] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2:246–252, 1999.

[15] T. Tanaka, A. Shimada, R. Taniguchi, T. Yamashita, and D. Arita. Towards robust object detection: integrated background modeling based on spatio-temporal features. In *Asian Conference on Computer Vision 2009*, 2009.

[16] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: Principle and Practice of Background Maintenance. *International Conference on Computer Vision*, pages 255–261, 1999.

[17] S. Yoshinaga, A. Shimada, H. Nagahara, and R. ichiro Taniguchi. Object Detection Using Local Difference Patterns. In *Asian Conference on Computer Vision 2010*, 2010.

[18] S. Zhang, H. Yao, and S. Liu. Dynamic background modeling and subtraction using spatio-temporal local binary patterns. In *15th IEEE International Conference on Image Processing*, pages 1556 –1559, 2008.

[19] Z. Zivkovic and F. van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27:773–780, 2006.