Titre: Title:	RECTGAUSS-Tex : block-based background subtraction
Auteurs: Authors:	Dorra Riahi, Pier-Luc St-Onge et Guillaume-Alexandre Bilodeau
Date:	2012
Type:	Rapport / Report
Référence: Citation:	Riahi, D., St-Onge, P.-L. & Bilodeau, G.-A. (2012). *RECTGAUSS-Tex : block-based background subtraction* (Rapport technique n° EPM-RT-2012-03).

EPM–RT–2012-03

RECTGAUSS-TEX:
BLOCK-BASED BACKGROUND SUBTRACTION

Dorra Riahi, Pier-Luc St-Onge,
Guillaume-Alexandre Bilodeau
Département de génie informatique et génie logiciel
École Polytechnique de Montréal

Mai 2012

Poly

ÉCOLE
POLYTECHNIQUE
MONTRÉAL

EPM-RT-2012-03

RECTGAUSS-Tex: Block-based Background Subtraction

Dorra Riahi
Pier-Luc St-Onge
Guillaume-Alexandre Bilodeau
Département de génie informatique et génie logiciel
École Polytechnique de Montréal

Mai 2012

Dépôt légal :
Bibliothèque nationale du Québec, 2012
Bibliothèque nationale du Canada, 2012

EPM-RT-2012-03
RECTGAUSS-Tex: Block-based Background Subtraction
par : Dorra Riahi, Pier-Luc St-Onge, Guillaume-Alexandre Bilodeau
Département de génie informatique et de génie logiciel
École Polytechnique de Montréal

RECTGAUSS-Tex: Block-based Background Subtraction

Dorra Riahi, Pier-Luc St-Onge and Guillaume-Alexandre Bilodeau

Department of Computer Engineering and Software Engineering
École Polytechnique de Montréal.

{dorra.riahi, pier-luc.st-onge, guillaume-alexandre.bilodeau@polymtl.ca}

Abstract

This paper presents an approach to background subtraction based on rectangular regions (blocks). The general principle is to successively divide the image into blocks and detect foreground pixels based on the color histogram and the variance between pixels of the blocks. Then, the classic Gaussian Mixture background subtraction method is applied to refine the detected foreground. Results show that this approach reduces false positives by filtering noise coming from small motion as it is based on groups of pixels instead of on individual pixels.

1. Introduction

Motion detection is a crucial task in many computer vision applications, such as robotics, video monitoring, and action recognition. Several approaches to motion detection are based on background subtraction. The fundamental principle of background subtraction is to build a background model of an empty scene, and then compare that model with the current image**.** The difference forms the moving objects**.** However, irrelevant pixels can be detected as foreground (shadow, image noise, dynamic scene element, etc.). Thus, a background subtraction method has to be able adapt to different conditions in a video sequence and to a changing background. Most background subtraction methods label pixels as background or foreground based on pixel by pixel decision such as Single Gaussian (SG) [4], Kernel Density Estimation (KDE), Temporal Median Filter, etc. [5]. Thus, these methods can be sensitive to noise and small perturbations [1].

In this paper, we present a block-based background subtraction method, RECTGAUSS-Tex, originally proposed in [2]. We have slightly modified the original method to automatically determine the best block size based on the image resolution. In this method, background modeling is done at different scales based on color histograms and the textural content of image blocks. Results show that this method reduces the number of false positives.
The paper is structured as follows. Section 2 describes the method. Section 3 reports and comments the results on the change detection challenge dataset, and section 4 concludes the paper.

2. Methodology

Background subtraction is performed in two stages:

1. Divide the image iteratively into rectangular regions (blocks), modeling each of them using a color histogram and a texture measure. Compare the blocks from the coarsest scale to the finest scale using the MDPA (Minimum Difference of Pair Assignments) histogram distance. This gives coarse foreground detection at the scale of the smallest block (Figure 1).

2. Apply Gaussian Mixture Method (GMM) to detect the foreground at the pixel level for each foreground block (Figure 2).

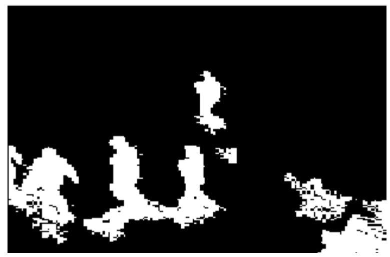

Figure 1. Background Subtraction with blocks only.

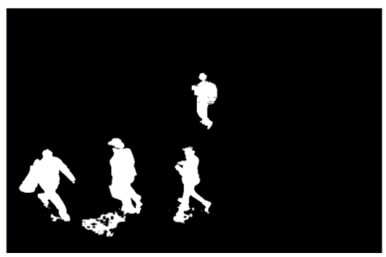

Figure 2. Background Subtraction after GMM.

2.1. Background modeling

The reference image (the first frame of a video sequence) is divided into blocks of size NxM. Originally, the blocks were 4x3, which is not always appropriate, except for 1.33 ratio images. Therefore, we made a change to the method of [2] in order to adapt to different image sizes. Thus, first, the reference image is divided into blocks of size NxM depending on the image ratio. For each block, a color histogram (64 bins for each RGB channel) and the variance of the pixels of the block are calculated. These two statistical measures captures the statistics of the pixels in the blocks and thus of the background. This is the finest scale. NxM blocks are then grouped together and their statistics are merged until a minimum number R_c (user defined) of blocks are obtained. Four blocks at the finest scale gives one block at the next scale, and so on. This gives background image M_R. The background is updated by substituting blocks that are labeled as background during motion detection.

2.2. Motion detection

To detect motion, a new frame I_R is modeled similarly using a hierarchy of blocks. The corresponding blocks in I_R and M_R at the coarsest scale are first compared. And only if they are different, blocks at a finer scale are compared and so on until the finest scale is reached (Figure 3). Blocks are compared using the color histograms and their variance. For each block of M_R and I_R, their histograms H_M and H_I, respectively, are compared using the MDPA distance S_H at position (i,j) with:

$$S_H\big(H_I\,(i,j), H_M\,(i,j)\big) = \frac{\sum_{b=0}^{K-1}\left|\sum_{k=0}^{b}\big(H_I\,(i,j)\big)[k] - \big(H_M\,(i,j)\big)[k]\right|}{\sum_{b=0}^{K-1}\big(H_I\,(i,j)\big)[b]} \quad (1)$$

Where b and k are the *k-th* and *b-th* histogram bin. The choice of MDPA is justified by the fact that this distance considers the error distribution among the histograms' bins, not only the sum of errors. Two histograms are similar if S_H is less than a threshold T (which is incremented by a Δt value for each scale). A similar process is used to compare the variances (texture information).

2.3. Integration of Gaussian Mixture Method (GMM)

After detection of the foreground using block comparison at the finest scale, improving the image foreground is needed because the detection is limited to small NxM blocks.

Consequently, the GMM method [6] is applied to the foreground blocks at the finest scale. Each pixel $p(x,y,t)$ is modeled by K distribution of Gaussian Mixture. The probability P that a pixel is a background pixel is defined by:

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} * \eta(X_t\,, \mu_{i,t}, \Sigma_{i,t}\, t) \; (2)$$

Where η is a Gaussian probability density function:

$$\eta(X_t, \mu_t, \Sigma_t\;) = \frac{1}{(2\pi)^{1/2}|\Sigma_t\;|^{1/2}}\, e^{-\frac{1}{2}(X_t-\mu_t)^T}\, \Sigma_t^{-1}(X_t - \mu_t) \;\; (3)$$

Where, X_t is a measure, μ_t is the mean and $\sum$ is the covariance of the distribution. A pixel X is modeled by K Gaussian distributions. For each one, we would choose the B best distributions. A pixel is part of the foreground if:

$$|X_t - \mu_t| > T_g \sigma_t$$

Where σ_t^2 is the variance and Tg is a threshold.

3. Experiments

3.1. Parameters choice

To test our method, we used the dataset of the change detection challenge [3]. Our method was tested for the following challenges: (Baseline, Camera Jitter, Intermittent Objects, Dynamic Background, Shadow and Thermal). In the experiment, we had to set the following parameters for all videos:

- R_C: the maximum number of merged blocks in the image. The sizes of each block corresponding to the image ratio (length / width).
- T: a similarity threshold between two color histograms. T varies according to the scale in order to avoid pixels noise. This variation depends on another parameter:ΔT. So, $T = T + \Delta T$ at each scale.

- A: learning rate for GMM.
- K: Number of Gaussian distributions.
- T_b: threshold for choosing the best B distributions.
- T_g: number of standard deviation.

To choose the best values for the parameters, we implemented an optimization algorithm. A range (minimum, maximum), a current value, and a change step are determined for each parameter. Thus, for each video, a set of combinations for different parameters is generated. In this method, GMM is applied only on blocks where motion is detected at the finest scale. Therefore, parameters optimization is first done only for the blocks, and then, an optimization for the entire method is deduced. After all the possible combinations of parameters were tested, the choice of a set of global parameters was performed using a selection from the best combinations for each video (a best combination is a combination with the best PBC and FNR). For the whole dataset, we used the parameters listed in table 1.

R_C	T	ΔT	A	K	T_b	T_g
80	0.022	0.0025	0.0045	5	0.9	2.5

Table 1. Parameters for the experiments using the detection challenge dataset [3].

3.2. Results

Different metrics are used to test the performance of our method. We used the metric suggested in the challenge, that is, precision, recall (Re), specificity (Sp), PBC, false negative rate (FNR), and false positive rate (FPR). Moreover, they depend on the number of true positive (TP), the number of false positive (FP), the number of false negative (FN) and the number of true negative (TN).

Method	Re	Sp	FPR	FNR	PBC	Precision
SOBS	0.9193	0.9987	0.0020	0.0026	0.4332	0.9313
GMM | KaewTraKulPong	0.8969	0.9980	0.0013	0.0194	1.9381	0.9532
ViBe	0.8872	0.9980	0.0020	0.0074	0.8869	0.9288
KDE	0.8385	0.9977	0.0023	0.0035	0.5499	0.9223
GMM | Stauffer & Grimson	0.8204	0.9972	0.0052	0.0112	1.5325	0.8461
GMM | Zivkovic	0.8180	0.9963	0.0028	0.0114	1.3298	0.8993
RECTGAUSS-Tex	0.6668	0.9979	0.002	0.0142	1.5341	0.9174
Mahalanobis distance	0.8085	0.9955	0.0037	0.0040	0.7290	0.9071
Euclidean distance	0.5863	0.9948	0.0045	0.0064	1.0260	0.9114

Table 2. Metrics for our method on the baseline dataset.

We present our result in tables that show our ranking for each category of images through a comparison of different metrics. As we can see in table 2 and table 6, in the baseline and thermal dataset, the false positive rate is (0.002 and 0.0006 respectively) which mean that most parts of the background are classified as background. However, our method has a smaller recall, because it may not detect small objects.

Despite the difficulty of the Camera Jitter dataset, our method is more successful compared to the baseline dataset (Table 3). Indeed, the recall metric is high (0.7648) because of the detection of small objects in the finer scale. Due to our two-step motion detection, most moving pixels from camera jitter are not detected as background. That is why the false negative rate is small (0.0092).

Method	Re	Sp	FPR	FNR	PBC	Precision
SOBS	0.8007	0.9787	0.0213	0.0075	2.7479	0.6399
GMM | KaewTraKulPong	0.5074	0.9888	0.0112	0.0205	3.0233	0.6897
ViBe	0.7112	0.9694	0.0306	0.0115	4.0150	0.5289
GMM | Stauffer & Grimson	0.7334	0.9666	0.0334	0.0109	4.2269	0.5126
KDE	0.7375	0.9562	0.0438	0.0101	5.1349	0.4862
RECTGAUSS-Tex	0.7648	0.9497	0.0502	0.0092	5.6662	0.4178
GMM | Zivkovic	0.6900	0.9665	0.0335	0.0127	4.4057	0.4872
Mahalanobis distance	0.7356	0.9431	0.0569	0.0105	6.4390	0.3813
Euclidean distance	0.7115	0.9456	0.0544	0.0115	6.2957	0.3753

Table 3. Metrics for our method on the Camera Jitter dataset.

Our method manages well occlusions between objects (intermittent object motion category) and the objects' shadows. Indeed, for tables 4 and 5, the value of a specificity is about 0.99 (rank 1). In addition, the rate of bad classification is relatively small (rank 2) for the intermittent Object videos. Thanks to the use of texture and region intensity variance, our method handles the shadows and small illumination changes.

Method	Re	Sp	FPR	FNR	PBC	Precision
SOBS	0.7057	0.9507	0.0493	0.0183	6.1324	0.5531
GMM | KaewTraKulPong	0.3476	0.9892	0.0108	0.0568	5.9854	0.6953
ViBe	0.5122	0.9527	0.0473	0.0425	7.7432	0.6515
KDE	0.5035	0.9309	0.0691	0.0466	10.0695	0.4609
RECTGAUSS-Tex	0.2189	0.9977	0.0022	0.0566	5.2546	0.5849
GMM | Stauffer & Grimson	0.5142	0.9835	0.0165	0.0421	5.1955	0.6688
GMM | Zivkovic	0.5467	0.9712	0.0288	0.0333	5.4986	0.6458
Mahalanobis distance	0.7165	0.8886	0.1114	0.0217	11.5341	0.4535
Euclidean distance	0.5919	0.9336	0.0664	0.0371	8.9975	0.4995

Table 4. Metrics for our method on the Intermittent Objects dataset.

Method	Re	Sp	FPR	FNR	PBC	Precision
SOBS	0.8350	0.9836	0.0164	0.0084	2.3366	0.7219
GMM | KaewTraKulPong	0.6323	0.9936	0.0064	0.0180	2.3015	0.8577
ViBe	0.7833	0.9919	0.0081	0.0094	1.6547	0.8342
KDE	0.8536	0.9885	0.0115	0.0063	1.6881	0.7660
RECTGAUSS-Tex	0.7026	0.9926	0.007	0.0145	2.067	0.8206
GMM | Stauffer & Grimson	0.7956	0.9871	0.0129	0.0103	2.2000	0.7156
GMM | Zivkovic	0.7770	0.9878	0.0122	0.0109	2.1957	0.7232
Mahalanobis distance	0.7845	0.9708	0.0292	0.0106	3.7896	0.5685
Euclidean distance	0.8001	0.9783	0.0217	0.0087	2.8987	0.6112

Table 5. Metrics for our method on the Shadow dataset.

Method	Re	Sp	FPR	FNR	PBC	Precision
SOBS	0.5888	0.9956	0.0044	0.0188	2.0983	0.8754
GMM | KaewTraKulPong	0.3395	0.9993	0.0007	0.0559	4.8419	0.9709
ViBe	0.5435	0.9962	0.0038	0.0323	3.1271	0.9363
KDE	0.6725	0.9955	0.0045	0.0140	1.6795	0.8974
GMM | Stauffer & Grimson	0.5691	0.9946	0.0054	0.0450	4.2642	0.8652
GMM | Zivkovic	0.5542	0.9942	0.0058	0.0451	4.3002	0.8706
RECTGAUSS-Tex	0.2461	0.9993	0.0006	0.0608	5.2656	0.9619
Mahalanobis distance	0.6270	0.9906	0.0094	0.0171	2.3462	0.8617
Euclidean distance	0.5111	0.9907	0.0093	0.0356	3.8516	0.8877

Table 6. Metrics for our method on the Thermal dataset.

Figure 3 presented some images resulting for processing the dataset videos. We can conclude that the best selection of parameters depends on the categories of the video, but still the method performs reasonably well for all scenarios.

The results show that our methods can filter some local noise, like trees or lake using the rectangular regions. In addition, our method deals with light shadows (shadow dataset) but not with large illumination changes.

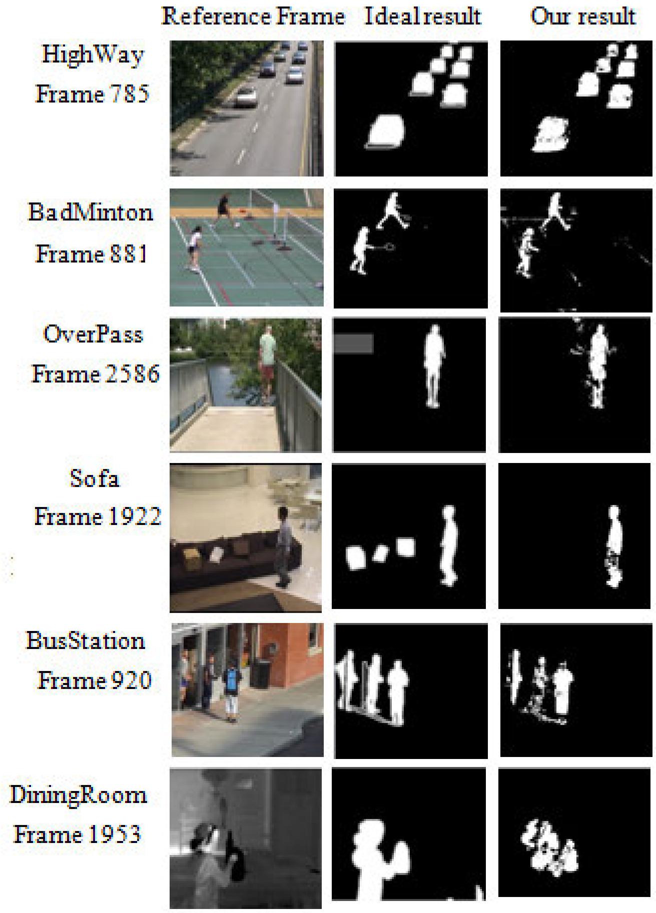

Figure 3. Detection results of our method for the dataset of the change detection challenge [3].

4. Conclusions

In this paper, the background subtraction method of [2] is applied on the dataset of [3]. This method is based on modeling the background with blocks at different scales. First, the background is modeled using blocks that are in turn modeled with a color histogram and the variance of intensities. Then, the Gaussian Mixture background subtraction method is applied to detect significant motion in the finest scale.

This method was evaluated in function of performance measures (FNR, Recall, etc.). Results show that our approach reduces false positives by filtering noise coming from small motion as it is based on groups of pixels instead of on individual pixels.

References

[1] M. Heikkila, M. Pietikainen, A Texture-Based Method for Modeling the Background and Detecting Moving Objects, IEEE Trans: Pattern Analysis and Machine Intelligence, vol. 28, no. 4, pp. 657-662, 2006.

[2] Darvish Zadeh Varcheie, P., Sills-Lavoie, M., Bilodeau, G.-A, A Multiscale Region-Based Motion Detection and Background Subtraction Algorithm, Sensors, 10 (2), 2010, pp. 1041-106.

[3] www.changedetection.net.

[4] McKenna, S., Jabri, S., Duric, Z., Rosenfeld, A., Wechsler, H, Tracking Groups of People. Computer Vision and Image Understanding, October 2000, vol. 80, pp. 42-56.

[5] Piccardi, M. , "Background subtraction techniques: a review," Systems, Man and Cybernetics, 2004 IEEE International Conference on , 10-13 Oct. 2004, vol. 4, pp. 3099- 3104.

[6] Turdu, D., Erdogan, H., "Improved post-processing for GMM based adaptive background modeling," Computer and information sciences, 2007. Iscis 2007. 22nd international symposium on, Nov. 2007, pp.1-6, 7-9.

L'École Polytechnique se spécialise dans la formation d'ingénieurs et la recherche en ingénierie depuis 1873

École Polytechnique de Montréal

École affiliée à l'Université de Montréal

Campus de l'Université de Montréal
C.P. 6079, succ. Centre-ville
Montréal (Québec)
Canada H3C 3A7

www.polymtl.ca

ÉCOLE
POLYTECHNIQUE
MONTRÉAL