Foreground Segmentation With Sudden Illumination Changes Using A Shading Model And A Gaussianity Test

Ka Ki Ng, Satyam Srivastava, Edward J. Delp *
Video and Image Processing Laboratory (VIPER)
School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana, USA

Abstract

In this paper, we propose a simple method for foreground segmentation based on a "Gaussianity test" and a shading model. The proposed method works under a hierarchical framework that combines a block based and a pixel based processing. The first step is a block-level classification based on the intensity differences and intensity ratios of a background model and the current frame. We then use pixel-wise adaptive background subtraction on the foreground classified blocks to obtain the foreground mask. We test our proposed method on several sequences with indoor scenes with extreme sudden illumination changes, and outdoor scenes under strong sunlight, waving tree leaves, and walking pedestrians. The method is shown to be robust to extreme sudden illumination changes and the presence of relatively small scene clutter motion (e.g. waving tree branches and leaves).

1 Introduction

In image analysis, foreground segmentation is the first step of many different image analysis applications, such as automated visual surveillance, video indexing, and human machine interaction. Since subsequent processes are greatly dependent on the performance of this step, it is important that the classified foreground pixels accurately correspond to the objects of interests. Many methods for moving object detection have been proposed. A typical approach to foreground segmentation is background subtraction which has very low computational cost. The goal here is to "remove" the background in a scene by describing an adequate model of the background with the result that only the "interesting" objects are left in the scene for tracking and further analysis. However, one drawback of traditional background subtraction is that it is vulnerable to environmental changes, for example, dynamic background (e.g., waving leaves), and gradual or sudden illumination changes.

Sudden illumination change is still a very challenging problem for foreground segmentation. Many state-of-the-art techniques can handle gradual illumination changes but remain susceptible to sudden changes. Some examples of sudden illumination changes are turning on/off light sources in a room, or open/close window curtains or doors. These situations cause color or intensity-based subtraction methods to fail (false positive i.e. detecting background pixels as foreground pixels).

One of the popular approaches, Gaussian Mixture Models (GMM) proposed by Stauffer and Grimson [1] is robust to gradual illumination as well as moving background regions. However, it is not robust to sudden illumination changes, foreground objects could be integrated into the background model if they remain static for a long period of time, and it has a relatively higher computational cost.

Vosters *et al.* [2] use Eigenbackground technique combined with a statistical illumination model for sudden illumination changes. Eigenbackground model is obtained from training data, and used to reconstruct the background image for each input image. Pilet *et al.* [3] propose a method that replaces the statistical background model by a statistical illumination model. They model the ratio of intensities between a stored background image and an input image in all three channels as a GMM to account for illumination effects. To take into account the dependence among neighboring pixels, a spatial likelihood model is obtained from ground truth data, and integrated into the framework. Javed *et al.* [4] use both color and gradient information in a pixel, region, and frame level framework. They use gradient based background subtraction and model the distribution of the gradient magnitudes and directions statistically to eliminate spurious objects and illumination changes. Xie *et al.* [5] suggest that the ordering among pixels is preserved even in the presence of heavy photometric distortions. He proposed a method using the Phong's shading model [6] with slowly spatially varying illumination, the sign of the difference between two pixel measurements is robust to sudden illumination changes. Nicolas and Pinel [7] proposed a sophisticated model for moving cast shadow segmentation

*This material is based upon work supported by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001. Address all correspondence to E. J. Delp (ace@ecn.purdue.edu).

(which is a local illumination change) and light source detection. Their method is based on the geometrical relations among light source, object and cast shadow under the assumption that the light source is unique and far form the scene. Their model is primarily used for shadow detection and not foreground segmentation. Some examples of block-based and descriptor-based background subtraction methods include [8] and [9]. In [8], Heikkila *et al.* use a texture feature that is based on local binary pattern (LBP) to describe background statistics of each block. Chen *et al.* [9] propose a novel feature called contrast histogram descriptor and integrate it into a hierarchical framework for block based and pixel based background modeling. Radke *et al.* [10] give an excellent survey on image change detection in the literature.

In this paper, we propose a simple method for foreground segmentation that is robust to sudden illumination change and small background dynamics. The method is an extension of a Gaussianity test proposed by Gurcan *et al.* [11, 12] for detection of microcalcifications in mammogram images. This test was developed by Ojeda *et al.* [13] for causal invertible time series data. The authors in [11, 12] make an assumption that when the microcalcification is absent, the difference image between the original image and the filtered image using a Least Mean Square adaptive filter are samples from a Gaussian distribution. If a region has a "Gaussianity test value" higher than a pre-determined threshold, then the region is classified as a region of microcalcification cluster. Otherwise, the Gaussianity test value (or *statistic*) should be close to zero.

Our proposed method is based on the assumption that the camera noise is *spatially* Gaussian, and we use the intensity differences of a background model and the current frame for the Gaussianity test. In the case of sudden illumination change, a shading model is imposed on the Gaussianity test for robustness. Instead of using the intensity differences, we use the intensity ratios (that is implied by the shading model) of a background model and the current frame for the Gaussianity test. We tested our proposed method on several sequences with indoor scenes under extreme sudden illumination changes, and outdoor scenes under strong sunlight, waving tree leaves, and walking pedestrians. Experimental results indicate our approach is robust to small/repetitive motions in dynamic scenes (such as waving tree leaves) and extreme sudden illumination changes.

2 Background Model Initialization

In order to avoid the assumption of a sequence starting in the absence of foreground objects, temporal frame differencing is used for the initial phase of the method until the background pixels are stable. Temporal frame differencing $FD_t(x, y)$ and background subtraction $D_t(x, y)$ are defined as:

$$FD_t(x, y) = |I_t(x, y) - I_{t-1}(x, y)|, \quad (1a)$$
$$D_t(x, y) = |I_t(x, y) - BM_{t-1}(x, y)|, \quad (1b)$$

where $I_t(x, y)$ is the intensity of pixel (x, y) in the current frame at time t, and $BM_t(x, y)$ is the intensity of pixel (x, y) of the background model at time t.

At the beginning of a sequence, only FD_t is computed because no background information is available. The FD_t is then compared to a threshold computed adaptively depending on the current frame as described in [14].

$$(x, y) \subset \begin{cases} foreground & \text{if } FD_t(x, y) > Th_{FD,ad} \\ background & \text{otherwise.} \end{cases} \quad (2)$$

If a pixel is determined to be a background pixel for Th_{frame} frames, it will be considered as a stable pixel and is used to construct the background model. As soon as background information is available for a pixel, $D_t(x, y)$ instead of $FD_t(x, y)$ is used:

$$(x, y) \subset \begin{cases} foreground & \text{if } D_t(x, y) > Th_{BS,ad} \\ background & \text{otherwise.} \end{cases} \quad (3)$$

As in Equation 2, the threshold $Th_{BS,ad}$ is obtained adaptively [14]. It is also assumed that every pixel of the background will be uncovered at some time.

3 Block Based Processing Using Gaussianity Test

A Gaussianity test determines if a set of samples are from a Gaussian distribution. The use of the test is inspired by Gurcan *et al.* [11, 12] who used it for detection of microcalcifications in mammogram images. The authors make an assumption that when the microcalcification is absent, the difference image between the original image and the filtered image using a Least Mean Square adaptive filter are samples from a Gaussian distribution. Our proposed method is based on the assumption that the camera noise is *spatially* Gaussian. In addition, we assume that this Gaussian noise is temporally uncorrelated and hence, independent across the frames. Hence only Gaussian noise and foreground objects remain in the difference frame $D_t(x, y)$ in Equation 1b (because the sum of independent Gaussian random variables is Gaussian). Under these assumptions, the foreground pixels in the difference frame $D_t(x, y)$ should be non-Gaussian distributed, and the background pixels in $D_t(x, y)$ should be Gaussian distributed.

We use the Gaussianity test in a block-based manner to detect non-Gaussianity. We consider a $M \times M$ block centered at pixel (x, y), and the test is based on the sample estimates of the first four moments of the pixel intensity differences given by:

$$\hat{J}_k(x, y) = \frac{1}{M^2} \sum_{m=-\frac{M-1}{2}}^{\frac{M-1}{2}} \sum_{n=1-\frac{M-1}{2}}^{\frac{M-1}{2}} [D_t(x + m, y + n)]^k \quad (4)$$

for $k = 1, 2, 3, 4$. In our experiments, we use $M = 17$ and 33 for each non-overlapping block in a frame. Note that in the rest of the paper, we will drop the pixel location (x, y)

when representing the first four moments $\hat{J}_1, \hat{J}_2, \hat{J}_3, \hat{J}_4$ for simplicity.

Consider the moment generating function of a Gaussian distribution with mean μ and σ^2:

$$M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}. \quad (5)$$

The k^{th} order moment of distribution J_k, is defined in terms of the moment generating function as:

$$J_k = E[D^k] = \frac{d^k}{dt^k} M(t)|_{t=0} \quad (6)$$

If a set of samples is Gaussian distributed, the sample moments converge to their theoretical values as the sample size goes to infinity under the ergodicity assumption, i.e.

$$\begin{aligned} J_1 &\rightarrow \mu \\ J_2 &\rightarrow \sigma^2 + \mu^2 \\ J_3 &\rightarrow \mu^3 + 3\sigma^2\mu \\ J_4 &\rightarrow \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 \end{aligned} \quad (7)$$

Similar to [11, 12] we construct our Gaussianity test by defining the Gaussianity test statistic as:

$$H(J_1, J_2, J_4) = J_4 + 2J_1^4 - 3J_2^2 \quad (8)$$

If we substitute the limit values in Equation (7) into Equation (8), then:

$$\begin{aligned} H(J_1, J_2, J_4) &= (\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4) + 2\mu^4 \\ &\quad -3(\sigma^2 + \mu^2)^2 \\ &= 0 \end{aligned} \quad (9)$$

Hence, the Gaussianity test statistic is expected to be close to zero if the set of samples is Gaussian distributed. In our experiments, if a set of samples in a block has a Gaussinity test statistic greater than a threshold τ, then the block is considered as a block that contains foreground pixels.

$$block = \begin{cases} \text{contains foreground pixels} & \text{if } H > \tau \\ \text{is background} & \text{otherwise.} \end{cases} \quad (10)$$

Note that this test can be used for other visual features for foreground segmentation under the assumption that the visual features are Gaussian (non-Gaussian) distributed in the absence (presence) of foreground objects.

4 Pixel Based Processing and Foreground Mask Generation

Once all blocks are classified as foreground/background blocks, pixel-based background subtraction is done on the foreground blocks to obtain a foreground object mask. The foreground mask is obtained by frame differencing or background subtraction depending on the availability of background information as described in Section 2. In other words, the segmentation mask is computed using Equation (2) or Equation (3) on the pixels contained *only* in the blocks classified as foreground blocks by the Gaussianity test. Morphological filtering is then done on the foreground segmentation mask to remove noise.

5 Sudden Illumination Change

Many state-of-the-art approaches [1, 15] can handle gradual illumination changes very well but remain vulnerable to sudden illumination changes. The assumption that the background regions are Gaussian distributed does not hold because there is photometric distortion in the case of sudden illumination changes. The increase of intensities is not a constant even when the illumination change is global. It results in a drastic increase in false positive detections. In many cases the entire image is detected as foreground. A shading model described below is imposed on the Gaussianity test to handle sudden illumination changes.

5.1 Shading model

Skifstad and Jain [16] propose an illumination-invariant method for change detection using a shading model. The model relates the intensities between the current frame $I_1(x, y)$ and a reference frame $I_2(x, y)$ in a given area of interest. A pixel intensity $I(x, y)$ is "decomposed" into a product of illumination $\pounds_i$ and a shading coefficient $S(x, y)$. The model is defined by:

$$I(x, y) = \pounds_i S(x, y). \quad (11)$$

The assumption of the model is that if there is no physical change (e.g. an object moving) between two frames, then the ratio expressed in Equation (12) is constant and independent of the shading coefficients $S(x, y)$:

$$R(x, y) = \frac{I_1(x, y)}{I_2(x, y)} = \frac{\pounds_{i,1}}{\pounds_{i,2}}, \quad (12)$$

In our proposed method, we determine this intensity ratio for each pixel of the current frame and the background model. It is assumed that when there is no foreground objects in the scene, the ratios of pixel intensities should be Gaussian distributed. We find the first four moments of the pixel intensity ratios $R(x, y)$ (similar to Equation 4):

$$J_k(x, y) = \frac{1}{M^2} \sum_{m=-\frac{M-1}{2}}^{\frac{M-1}{2}} \sum_{n=1-\frac{M-1}{2}}^{\frac{M-1}{2}} [R(x+m, y+n)]^k, \quad (13)$$

and the Gaussianity test described by Equations (8) and (10) as an illumination-invariant foreground segmentation method.

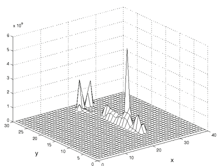
6 Experimental Results

In our experiments, we use test sequences with indoor scenes under extreme sudden illumination changes, and outdoor scenes under strong sunlight, waving tree leaves, and walking pedestrians. These are considered the challenging cases for foreground segmentation. We also test the proposed method on a publicly available dataset PETS 2006 [17].

Results of the block based detection result is shown in Figure 1. Figures (a,e) and (b,f) are the background frame and the current frame respectively, Figure (c,g) is a 3D plot of $H(x, y)$ using Equation (8), and Figure (d,h) shows the foreground pixels determined by Equation (10) in white.

(a) Background frame.

(b) Current frame.

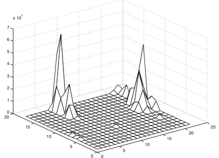


(c) 3D plot of $H(x, y)$.

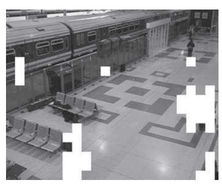
(d) Detection results.

(e) Background frames.

(f) Current frames.



(g) 3D plots of $H(x, y)$.



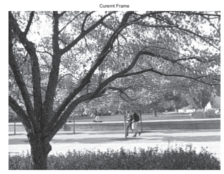
(h) Detection results.

Figure 1. Block based detection results. (a,e) Background frame. (b,f) Current frame. (c,g) The H values of the foreground blocks are much higher than the background blocks. (d,h) The blocks detected as foreground are white.

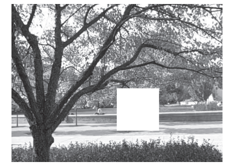
The test can distinguish background from foreground blocks accurately. The threshold $\tau = 10^5$ is empirically determined – as shown in Figure 6, the H values of the foreground blocks are very discriminative and are much higher than the H values of the background blocks.

Figure 2 shows the results for an outdoor scene with a dynamic background. It is observed that the distribution of the intensity difference of scenes with waving tree leaves resembles a Gaussian distribution closely. It has false alarms for the block-level processing (i.e. blocks are non-Gaussian in leaves region), however the pixel-level processing (i.e. background subtraction or temporal frame differencing) is still able to obtain a fine foreground segmentation mask.

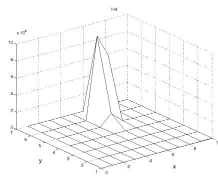
Figures 3 and 4 show the results of processing two sequences under extreme sudden illumination change, using the shading model. They both have grayscale intensity changes of more than 100. The sequence in Figure 4 is a very challenging test sequence. A person appears in the scene (as shown in Figure 4b) but it was so dark that even humans can barely notice the person. Both GMM and our proposed method *without* the shading model fail and detect many false positives (almost the entire frame) in this case. However, using our proposed method with shading model, the number of false negatives decreases drastically.



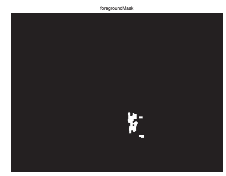
(a) Waving leaves sequence.



(b) Gaussian distributed blocks.

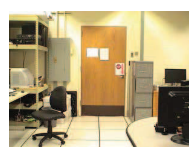


(c) 3D plot of $H(x, y)$.

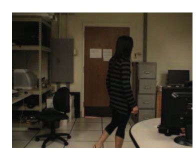


(d) Fine foreground mask.

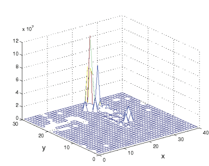
Figure 2. (a) Current frame with waving tree leaves. (b) The Gaussian distributed blocks (foreground) detected are white. (c) 3D plot of $H(x, y)$. (d) Fine foreground segmentation mask after pixel-level processing.



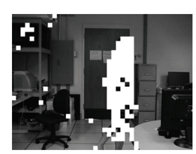
(a) Background model .



(b) Current frame.



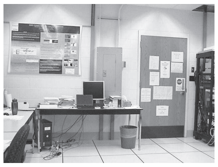
(c) 3D plot of $H(x, y)$.



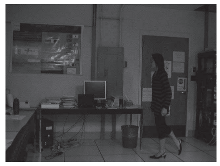
(d) Detection results.

Figure 3. (a) Background model with average grayscale intensity = 156. (b) Current frame with average grayscale intensity = 55. (c) 3D plot of $H(x, y)$. (d) The blocks detected as foreground are white.

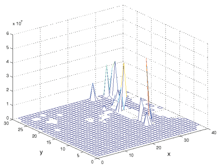
We compare our method with the popular GMM tech-



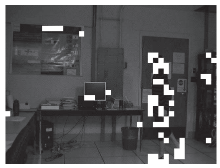
(a) Background model .



(b) Current frame.

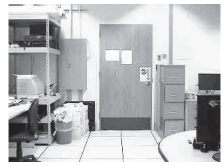


(c) 3D plot of $H(x, y)$.



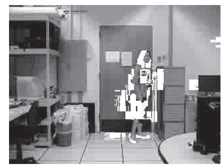
(d) SM detection results.

Figure 4. (a) Background model with average grayscale intensity = 157. (b) Current frame with average grayscale intensity = 32. Note that there is a person in front of the door. (c) 3D plot of $H(x, y)$. (d) The blocks detected as foreground are white.



(a) Background model .

(b) Current frame.



(c) Our proposed method.

(d) GMM.

Figure 5. (a) Background model with average grayscale intensity = 215. (b) Current frame with average grayscale intensity = 189. Note that the intensity difference is only 26 in this case. (c) Our proposed method detect part of the foreground. (d) The white area is detected as foreground, GMM fails in this case.

nique using a test sequence with sudden illumination change. The sequence is taken indoors with a person switching off some of the overhead lights. The average grayscale intensity change in this case is moderate. It decreases from 215 to 189, a change of only 26 (compared to over 100 in the previous two experiments). Figure 5 shows the results of the comparison. GMM fails and detects many false positives (almost the entire frame) in this case. The reason is that the pixels that have large changes in intensities are detected as foreground because they do not match any background distribution in the Gaussian mixture. Using our proposed method, majority of the foreground object is detected.

7 Conclusions and Future Work

In this paper, we propose a simple method for foreground segmentation using a shading model and a Gaussianity test. Our background model initialization technique is robust to the presence of foreground objects. Our proposed segmentation method consists of a block-level and pixel-level processing based on the intensity differences and ratios of the background model and the current frame. Experimental results illustrate our approach is robust to small or repetitive motions in dynamic scenes such as waving tree leaves, and to extreme sudden illumination changes. Future extensions of this work include the use of ground truth for quantitative results, an investigation of how the block size of the block-level processing affects the performance of the Gaussianity test, and a computational cost analysis of our proposed method.

References

[1] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, August 2000.

[2] L. Vosters, C. Shan, and T. Gritti, “Background subtraction under sudden illumination changes,” *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance*. Boston, Massachusetts: IEEE Computer Society, August-September 2010, pp. 384–391.

[3] J. Pilet, C. Strecha, and P. Fua, “Making background subtraction robust to sudden illumination changes,” *Proceedings of the European Conference on Computer Vision*, Marseille, France, October 2008, pp. 567–580.

[4] O. Javed, K. Shafique, and M. Shah, “A hierarchical approach to robust background subtraction using color and gradient information,” *Proceedings of the IEEE Workshop on Motion and Video Computing*, Orlando, Florida, December 2002, pp. 22–27.

[5] B. Xie, V. Ramesh, and T. Boult, “Sudden illumination change detection using order consistency,” *Image and Vision Computing*, vol. 22, no. 2, pp. 117–125, February 2004.

[6] B. T. Phong, “Illumination for computer generated pictures,” *Communications of the ACM*, vol. 18, no. 6, pp. 311–317, 1975.

[7] H. Nicolas and J.-M. Pinel, "Joint moving cast shadows segmentation and light source detection in video sequences," *Signal Processing: Image Communication*, vol. 21, no. 1, pp. 22–43, 2006.

[8] M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 657–662, 2006.

[9] Y.-T. Chen, C.-S. Chen, C.-R. Huang, and Y.-P. Hung, "Efficient hierarchical method for background subtraction," *Pattern Recognition*, vol. 40, no. 10, pp. 2706 – 2715, 2007.

[10] R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 294–307, March 2005.

[11] M. N. Gurcan, Y. Yardimci, and A. E. Cetin, "Influence function based gaussianity tests for detection of microcalcifications in mammogram images," *Proceedings of the IEEE International Conference on Image Processing*, vol. 3, Kobe, Japan, October 1999, pp. 407–411.

[12] M. N. Gurcan, Y. C. Yardimci, and E. A. Cetin, "2d adaptive prediction-based gaussianity tests in microcalcification detection," *Proceedings of the SPIE/IS&T Visual Communications and Image Processing*, vol. 3309, San Jose, CA, January 1998, pp. 625–633.

[13] R. Ojeda, J.-F. Cardoso, and E. Moulines, "Asymptotically invariant gaussianity test for causal invertible time series," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, Munich, Germany, April 1997, pp. 3713–3716.

[14] K. K. Ng and E. J. Delp, "Object tracking initialization using automatic moving object detection," *Proceedings of the SPIE/IS&T Visual Information Processing and Communication*, vol. 7543, San Jose, CA, January 2010.

[15] A. Elgammal, R. Duraiswami, D. Harwood, and L. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151–1163, July 2002.

[16] K. Skifstad and R. Jain, "Illumination independent change detection for real world image sequences," *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 3, pp. 387–399, June 1989.

[17] "Performance Evaluation of Tracking and Surveillance Dataset 2006," Public dataset http://www.cvg.reading.ac.uk/PETS2006/, Accessed: April 2011.