

# PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

## Small moving object detection using adaptive subband decomposition in video sequences

Rabi Zaibi, Enis A. Cetin, Yasemin C. Yardimci

**SPIE.**

# Small Moving Object Detection Using Adaptive Subband Decomposition in Video Sequences

Rabi Zaibi, A. Enis Çetin

Department of Electrical and Electronics Engineering,  
Bilkent University,  
Ankara, Turkey  
E-mail:cetin@ee.bilkent.edu.tr

Yasemin Yardımcı

Informatics Institute  
Middle East Technical University,  
Ankara, Turkey  
E-mail:yardimy@ii.metu.edu.tr

## ABSTRACT

In this paper, a small moving object method detection method in video sequences is described. In the first step, the camera motion is eliminated using motion compensation. An adaptive subband decomposition structure is then used to analyze the motion compensated image. In the “low-high” and “high-low” subimages small moving objects appear as outliers and they are detected using a statistical Gaussianity detection test based on higher order statistics. It turns out that in general, the distribution of the residual error image pixels is almost Gaussian. On the other hand, the distribution of the pixels in the residual image deviates from Gaussianity in the existence of outliers. Simulation examples are presented.

**Keywords:** Moving object detection, adaptive subband decomposition, wavelet transform, higher order statistics.

## 1. INTRODUCTION

In this paper, a small moving object method detection method in video sequences based on adaptive subband decomposition and higher order statistics is described. Detection of small moving objects can be a complicated task when there is noise and the video camera is in motion. Classical object detection methods [1], [2] are geared for large objects with clear features and boundaries whereas in our problem the moving region or object may consist of only a few pixels.

In our method, the first step is the elimination of the camera motion using motion compensation. After motion compensation, the resulting image basically contains the moving regions and objects. This image is processed using a two-dimensional (2-D) adaptive filter bank [3] in which the filters are updated according to an Least Mean Square (LMS) type adaptation algorithm. In this filterbank structure, each pixel is adaptively predicted using an appropriate neighborhood structure and four subimages are obtained. It turns out that the distribution of the “low-high” and “high-low” subimage pixels is almost Gaussian in general. However, small moving objects produce outliers in the residual image as the pixels of the small moving objects cannot be predicted accurately using the neighboring pixels. We detect the outliers using Higher Order Statistical (HOS) based Gaussianity test [8], [9]. In static regions the test statistics is very close to zero whereas in regions containing the moving object(s), the distribution of pixels deviate from Gaussianity and the test statistic produces large values.

Here is the outline of this paper. In Section 2, we present the 2-D adaptive subband decomposition method which removes the static background. In Section 3, we review the Higher Order Statistics (HOS) based Gaussianity test for moving object detection and present the results of simulation studies in Section 4.

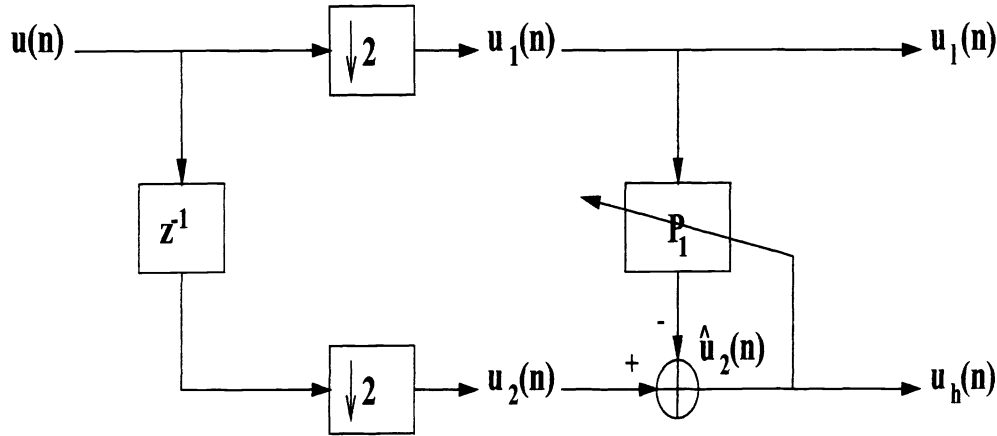


Figure 1. Adaptive subband decomposition structure.

## 2. ADAPTIVE SUBBAND DECOMPOSITION

The concept of adaptive subband decomposition is developed in [3, 4]. Adaptive subband decomposition can be considered as a trade-off between the adaptive prediction and ordinary lifting [11] based wavelet transform.

The adaptive subband decomposition structure [3]- [6] is shown in Figure 1. The structure was developed for one-dimensional signals, but we can apply it to two-dimensional signals by using the row by row and column by column filtering methods as in 2-D separable subband decomposition (or wavelet transform).

The first subsignal  $u_l$  is a downsampled version of the original signal  $u$ , a one dimensional signal which is usually a column or a row of the input image. As  $u_l$  is the result of a down-sampling by 2 operation, it contains only the even samples of the signal  $u$ . The sequence  $u_2$  is a shifted and downsampled by 2 version of  $u$ , containing only odd samples of  $u$ . We predict  $u_2$  using  $u_1$  and subtract the estimate of  $u_1$  from  $u_2$  to obtain the signal  $u_h$  which contains unpredictable regions such as edges of the original signal.

Various adaptation schemes can be used for the predictor  $P_1$ . In our work, we used the adaptive FIR estimator, as it proved to be good for the sample images that have been tested. This adaptive FIR estimator is obtained by predicting the odd samples  $u_2(n)$  from the even samples  $u_1(n)$  as follows:

$$\hat{u}_2(n) = \sum_{k=-N}^N w_{n,k} u_1(n-k) = \sum_{k=-N}^N w_{n,k} u(2n-2k) \quad (1)$$

The filter coefficients  $w_{n,k}$ 's are updated using an LMS-type algorithm [12] as follows:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \frac{\tilde{\mathbf{v}}_n e(n)}{\|\tilde{\mathbf{v}}_n\|^2} \quad (2)$$

where  $\hat{\mathbf{w}}(n) = [w_{n,-N}, \dots, w_{n,N}]$  is the weight vector at time instant  $n$ ,

$$\tilde{\mathbf{v}}_n = [u_1(n-N), u_1(n-N+1), \dots, u_1(n+N-1), u_1(n+N)]^T, \quad (3)$$

The subsignal  $u_h$  is given by

$$u_h(n) = u_2(n) - \hat{u}_2(n). \quad (4)$$

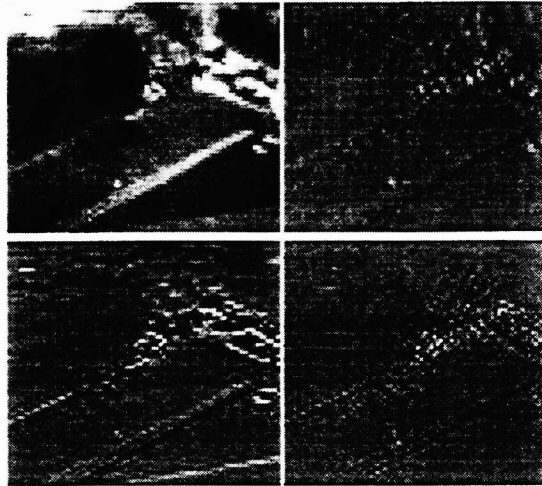
where  $u_h$  is the error we make in predicting the odd samples from the even samples, thus,

$$e(n) = u_h(n) = u_2(n) - \tilde{\mathbf{v}}_n^T(n) \hat{\mathbf{w}}(n). \quad (5)$$

Both  $\ell^1$  and  $\ell^2$  norms can be used in normalizing the update equation in (2) depending on the characteristics of the signal [12]. In this paper, the regular Euclidian norm is used. For the initial filter one can



**Figure 2.** An image from a video showing a car leaving a parking lot



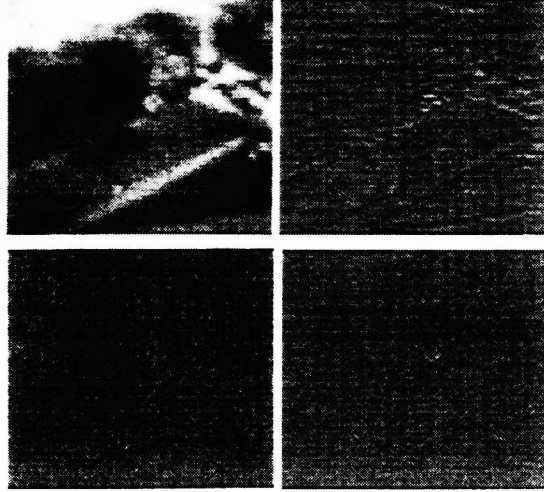
**Figure 3.** Subimages obtained using adaptive subband decomposition: Low-low, low-high and high-low, and high-high subimages.

use a typical lowpass filter for the adaptive predictor. The convergence of the adaptive filter is observed to be fast in natural images.

This structure is the simplest adaptive filterbank. Other adaptive filterbanks in which the “low-band” subsignal is a lowpass filtered and downsampled version of the original signal can be found in [3].

If the motion compensated image is processed by an adaptive filterbank we expect that small moving objects cannot be predicted as good as the other regions. Thus outliers will appear in  $u_h[n]$  in regions corresponding to moving objects.

The extension of the adaptive filterbank structure to two dimensions is straightforward. As in the case of ordinary subband decomposition, we process the image rowwise first and obtain two subimages. Consequently, these two subimages are processed columnwise and four subimages  $x_{ll}$ ,  $x_{lh}$ ,  $x_{hl}$ , and  $x_{hh}$  are obtained. Figure 2 shows the original image  $x$ , and the resulting subimages  $x_{ll}$ ,  $x_{lh}$ ,  $x_{hl}$  and  $x_{hh}$  obtained after adaptive subband decomposition are shown in Figure 3, respectively.



**Figure 4.** Subimages obtained using subband decomposition: Low-low, low-high and high-low, and high-high subimages.

This image is also processed by an ordinary wavelet transform. The resulting subimages are shown in Figure 4. The ‘low-high’ and ‘high-low’ images are sharper and the edges of the objects are highlighted more in the adaptive subband case. Adaptive subband decomposition gives better results in moving target detection for this reason.

### 3. HIGHER ORDER STATISTICAL TEST

It is experimentally observed that in regions with no moving objects, the subimages  $x_{lh}[m, n]$  and  $x_{hl}[m, n]$  have Gaussian like distribution, whereas in regions containing small moving objects it contains outliers. This is due to the fact that pixels of a small moving object cannot be accurately predicted using the surrounding pixels.

Higher order statistical tests are successfully used in the detection of microcalcifications in mammogram images [9], and detecting objects in noisy images [10]. In this paper, we use a Gaussianity test developed in [9] and [8]. The higher order statistic  $h(I_1, I_2, I_3, I_4)$  is based on the sample estimates of the first four moments  $I_1, I_2, I_3, I_4$  of the prediction error [8]. Estimates of the moments are given by

$$I_k = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N e^k[m, n] \quad (6)$$

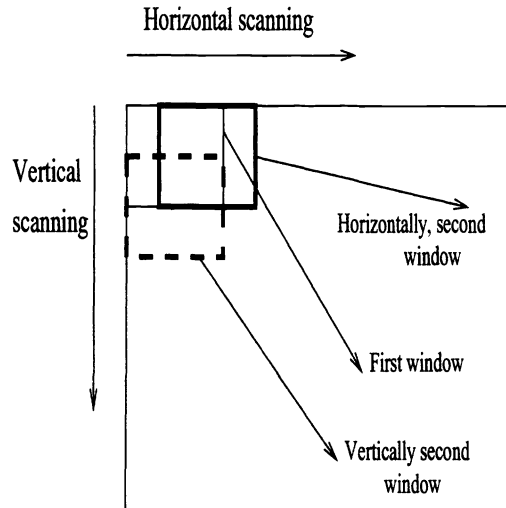
where  $e[m, n]$  represents the sum of the pixel values  $x_{lh}[m, n]$  and  $x_{hl}[m, n]$  and  $M \times N$  is the size of the region in which  $I_k$  is estimated. The subimages,  $x_{lh}$  and  $x_{hl}$  are obtained by processing the motion compensated image using the adaptive subband decomposition. The statistic  $h(I_1, I_2, I_3, I_4)$  is defined as follows:

$$h(I_1, I_2, I_3, I_4) = I_3 + I_4 - 3I_1(I_2 - I_1^2) - 3I_2^2 - I_1^3 + 2I_1^4 \quad (7)$$

It is ideally equal to zero when the distribution is Gaussian. It takes large values when the underlying distribution deviates from Gaussianity. Outliers in the error image are mainly due to moving objects and  $h(I_1, I_2, I_3, I_4)$  takes large values in such regions containing outliers.

In our approach, the video containing a moving target(s) is analyzed as follows:

- A motion compensated image is obtained from two consecutive images,
- adaptive subband decomposition of the motion compensated image is computed,



**Figure 5.** Higher order statistic computation in overlapping windows.

- the resulting subimages  $x_{lh}[m, n]$  and  $x_{hl}[m, n]$  are summed (the high-high subimage  $x_{hh}[m, n]$  contains almost no information for most practical images and it is not used in our algorithm) and analyzed block by block.

The HOS based statistic (7) is calculated within each block inside the image. These blocks may overlap as shown in Figure 5. In our experimental work we used blocks of size  $M = 15$  by  $N = 15$  where overlapping occurs at 3 pixel steps. If the HOS based statistic exceeds a threshold value in a block then this block is marked as a region containing a moving object. The above procedure is carried out over the entire video sequence.

As described above in each image block a statistical test is carried out to detect the moving object(s). The detection procedure can be considered as a hypothesis testing problem in which the null hypothesis  $H_0$  corresponds to the no moving object case and  $H_1$  corresponds to the presence of a moving object:

- $H_0: |h(I_1, I_2, I_3, I_4)| < T_h$
- $H_1: |h(I_1, I_2, I_3, I_4)| \geq T_h$

The threshold  $T_h$  is experimentally determined. The blocks in which the test statistic exceeds the threshold,  $T_h$ , are marked as regions containing the small moving objects.

#### 4. EXPERIMENTAL RESULTS

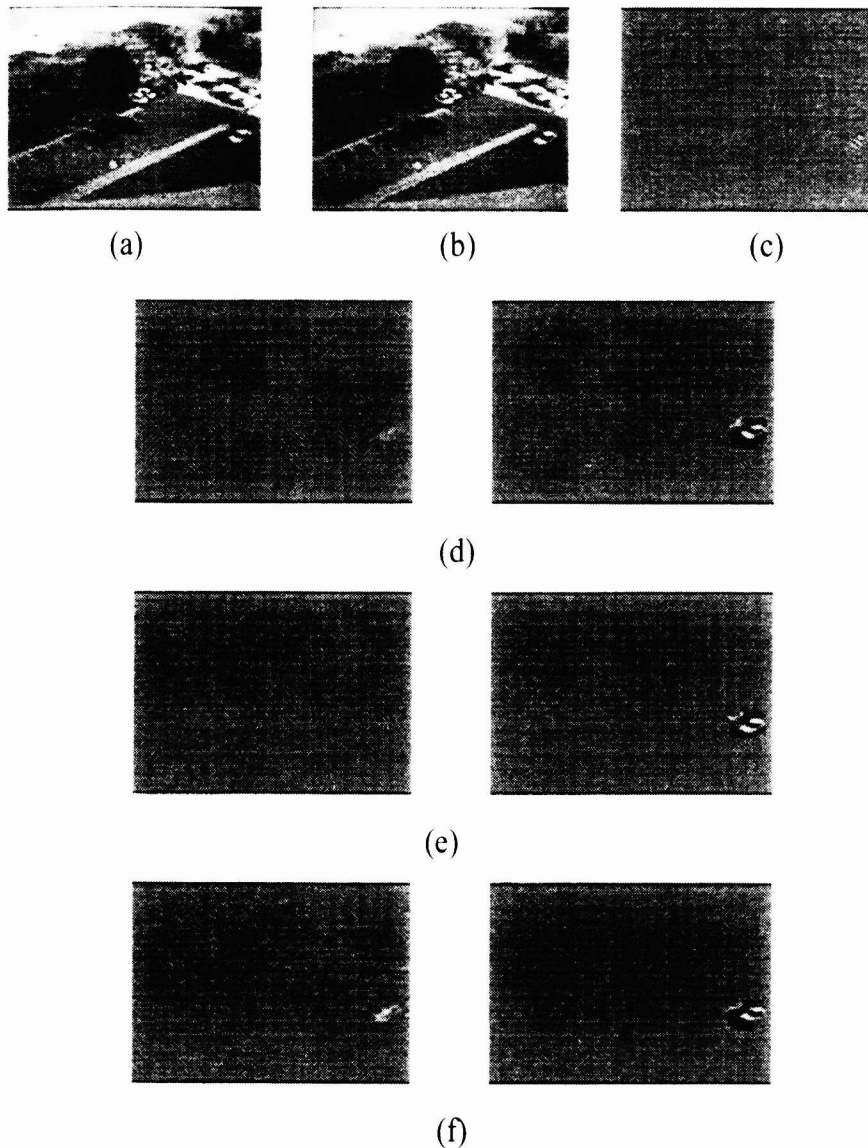
In this section, we present simulation studies. We test the performance of the detection scheme by analyzing 27 video sequences containing small moving objects on various backgrounds. As described in Section 1 and 2, motion compensated images are obtained in the first step. A classical block matching based motion compensation algorithm with subpixel accuracy is used [2].

In the second step, motion compensated images are filtered using the adaptive wavelet transformer and the subimages  $x_{lh}[m, n]$  and  $x_{hl}[m, n]$  are obtained. Finally, the test statistic values are obtained in small overlapping blocks. The values of the test statistic  $h$  in 12 video sequences are given in Table 1. It is clear from this table that a threshold can be selected which can distinguish moving objects from the background.

In our detection scheme we use an adaptive threshold value which is determined from the first image of the video sequence according to the following formula:

$$T_h = \frac{A}{L} \sum_m h_m(I_1, I_2, I_3, I_4) \quad (8)$$

where  $h_m(I_1, I_2, I_3, I_4)$  represents the value of the statistic  $h$  inside the  $m^{th}$  block, and  $L$  represents the total number of blocks inside the image. In our case  $A$  is chosen as 12 so that  $T_h$  is well above the maximum  $h$



**Figure 6.** (a) and (b) Two consecutive image frames from a video in which a car is leaving a parking lot, (c) motion compensated image, (d) adaptive prediction error image and the detected region (right) [13], (e) sum of the subimages  $x_{lh}[m, n]$  and  $x_{hl}[m, n]$  obtained after ordinary wavelet transform and the detected region (right), (f) sum of the subimages  $x_{lh}[m, n]$  and  $x_{hl}[m, n]$  obtained after adaptive subband decomposition and the detected region (right).

value of the blocks containing no moving target (maximum  $h$  value is 0.5 in the training set of 12 videos). Detection results are summarized in Table 2. In all of the 15 test videos different from the 12 training videos moving objects are determined successfully.

We also compared the performance of the adaptive predictor to the wavelet transform, and adaptive subband decomposition [3]. Motion compensated images are analyzed using (i) the adaptive predictor described in [8,13], (ii) wavelet transform (subband decomposition), and (iii) adaptive subband (or adaptive wavelet) decomposition [3].

Regions	Minimum	Maximum
With moving object	2.3	7.5
Without moving object	-0.41	0.5

**Table 1.** Values of the test statistic  $h(I_1, I_2, I_3, I_4)$  in regions with and without moving objects.

Algorithms	False Alarms	Miss
Adaptive prediction	0	0
Adaptive wavelet	2	0
Wavelet transform	0	4

**Table 2.** Detection performance of each method.

Typical results of the above methods are shown in Figure 4. The detection performance of these methods are summarized in Table 2. In the test videos, the adaptive predictor produces the best results. Adaptive subband decomposition also detects all of the moving objects but, in two cases, it produces false alarms. In ordinary wavelet transform, 4 targets are missed. By reducing the threshold all of the targets can be detected but in this case, the number of false alarms drastically increases.

## 5. CONCLUSION

In this paper, a small moving target detection method is proposed. The method is based on adaptive subband decomposition and higher order statistics. Experimental results indicate that the proposed method is effective and computationally efficient.

Adaptive subband decomposition provides a good trade-off between adaptive prediction and the ordinary wavelet transform in terms of detection performance and the computational cost.

The computational cost of the adaptive prediction based method [13] is much higher than the adaptive subband decomposition based method in which a quarter size image  $x_{lh} + x_{hl}$  is analyzed. Whereas in adaptive prediction based method HOS test computations are carried out over the entire image  $x$ .

## REFERENCES

1. A. A. Alatan, L. Onural, M. Wollborn, R. Mech, E. Tuncel, and T. Sikora. "Image sequence analysis for emerging interactive multimedia services - the European COST 211 framework," *IEEE Trans. on CAS for Video Tech.*, Nov. 1998.
2. A. M. Tekalp, *Digital Video Processing*. Prentice-Hall, 1995.
3. Ö. N. Gerek, A. E. Çetin, "Linear/nonlinear adaptive polyphase subband decomposition structure for image compression" *Proc. of IEEE ICASSP'98*, Seattle, U.S.A, May 1998; also accepted for publication in *IEEE Trans. Image Processing*.
4. Omer N. Gerek, A. E. Cetin, 'Polyphase adaptive filter banks for fingerprint image compression,' *Electronics Letters*, pp. 1931-1932, vol. 34, Oct. 1998.
5. Omer Nezh Gerek, A. Enis Cetin, "Polyphase Adaptive Filter Banks for Fingerprint Image Compression" *EURASIP Int'l. Conf. EUSIPCO'98*, Sept. 1998.
6. R.Oktem, K. Egiazarian, A. E. Cetin, "Subband Decomposition Based Image Compression Algorithms With Nonlinear Adaptive Filter Banks", *Proc. of IEEE- EURASIP NSIP 99*, Antalya, Turkey, vol. 2, pp. 766-769, June 1999.
7. R. Rajagopalan, E. Feig, and M. T. Orchard, "Motion optimization of ordered blocks for overlapped block motion compensation". *IEEE Trans. on CAS for Video Tech.*, Apr 1998.
8. M. N. Gürçan, Y. Yardımcı and A. E. Çetin, "Influence function based Gaussianity tests for detection of microcalcifications in mammogram images ", *IEEE International Conference on Image Processing (ICIP'99)*, Kobe, Japan, Oct. 1999.
9. R. Ojeda, J. Cardoso, and E. Moulines, "Asymptotically Invariant Gaussianity Test for Causal invertible Time Series", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97)*, vol. 5, pp 3713-3716, April 21-24, 1997.
10. G. B. Giannakis and M. K. Tsatsanis, "Time domain tests for Gaussianity and time-reversibility", *IEEE Trans. on Signal Processing*, vol. 42, pp. 3460-3472, 1994.



11. W. Sweldens. "The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Constructions", in *Proc. of Society of Photo-Optical Instrumentation Engineers (SPIE)*, vol. 2569, pp. 68-79, Sept. 1995.
12. O. Arikan, A. E. Çetin, Engin Erzin, "Adaptive Filtering for non-Gaussian stable processes", *IEEE Signal Processing Letters*, vol. 1, No. 11, pages 163-165, Nov. 1994.
13. Rabi Zaibi, Yasemin Yardimci, A. Enis Cetin, "Small Moving Object Detection In Video Sequences," to be presented in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2000)*, Istanbul, Turkey, June 2000.