ADAPTIVE THRESHOLDING FOR DETECTION OF NONSIGNIFICANT VECTORS IN NOISY IMAGE SEQUENCES

Luc Martel and André Zaccarin

Department of electrical and computer engineering
Laval University, Ste-Foy, Québec, Canada G1K-7P4
Email: martel@gel.ulaval.ca zaccarin@gel.ulaval.ca

ABSTRACT

In noisy image sequences, block matching motion estimation generates erroneous motion vectors since the algorithm tries to correlate noise. We present an adaptive threshold test to detect blocks for which only nonsignificant motion vectors can be estimated. Vectors of these blocks are then assigned the zero vector before any block motion estimation is performed. By nonsignificant, we refer to motion vectors of non moving areas as well as vectors of moving areas for which the noise level is too high to allow a good estimation of the motion. The detection of these vectors reduces the computational complexity of the BMA and the entropy of the motion field. The algorithm is embedded in a hierarchical BMA and takes advantage of their different spectral characteristics to discriminate between the frame difference energy due to noise and due to motion. The algorithm is also efficient for low noise sequences where it can be used to initialize a segmentation of moving objects from the background.

1. INTRODUCTION

Videoconferencing, movies and television are the standard applications for video compression. For all these media, the sequences to be processed are typically of good quality, with low noise content. For other less traditional media, however, quality of the image may not be as good, *e.g.*, sonograms in medical applications or infrared imagery for surveillance purpose. Our application deals with a surveillance scenario for which infrared images have to be compressed while allowing for the detection and tracking of dim targets.

In noisy image sequences, motion estimation by block matching algorithm (BMA) will often generate erroneous motion vectors. Computational resources are spent trying to correlate noise, and as a result, motion vectors often carry little or no information. Our objective is to detect blocks for which only nonsignificant vectors can be estimated, and set their motion vectors to (0,0). By nonsignificant, we refer to motion vectors of non moving areas as well as vectors of moving areas for which the noise level is too high to allow a good estimation of the motion. The detection of these vectors will reduce the computational complexity of the BMA and the entropy of the motion field.

One way to detect moving regions is to threshold the difference between two consecutive frames, as is done by Diehl in [4]. Pixels with an absolute value above a certain level are considered as moving and the others stationary. A block can then be considered as stationary if the number of its stationary pixels is higher than a given number. This procedure is described by Salari and Lin in [1]. A block can also be labeled as stationary if the energy of the frame difference is below a given threshold [5]. In the MPEG guide to coding [6], it is also suggested to use the zero motion vector if its prediction error is slightly higher than that of another vector because of possible coding gains. Such a strategy, however, does not offer any savings in computation time.

All these methods require we choose one or several threshold values. We present in this paper an adaptive threshold test for detection of nonsignificant vectors. The test uses an estimation of the noise power of the image sequence and is embedded in a hierarchical block matching algorithm (HBMA). The threshold test is applied to low pass filtered images used by the HBMA. Because of the different spectral characteristic of the frame difference caused by noise and motion, we will see that the low pass filter allows us to adapt the threshold value of the test.

The paper is organized as follows. We present the image model and the noise power estimation in the following section. The proposed adaptive threshold test, and its motivation, are presented in section 3. Results and conclusions follow in sections 4 and 5.

2. NOISE POWER ESTIMATION

2.1. Image model

For the purpose of computing an estimate of the noise power, the n^{th} image of the sequence, $I_n[x,y]$, is modeled as

$$I_n[x, y] = S_n[x, y] + N_n[x, y] \tag{1}$$

where the signal $S_n[x, y]$ is a random field of mean μ_S and variance σ^2_S, and where the noise $N_n[x,y]$ is a zero mean white random field of variance σ^2_N. We further assume that there is no correlation between the noise and the signal, and that there is also no correlation between the random variables $N_{n-1}[x,y]$ and $N_n[x,y]$. Under these hypotheses, the variance of the difference between two consecutive frames is given by

$$Var\{D[x, y]\} = 2\sigma_N^2 + 2\sigma_S^2(1-r) \tag{2}$$

where $D[x, y] = I_{n-1}[x, y] - I_n[x, y]$, and where r is the is the correlation coefficient between S_{n-1} and S_n.

If there is no motion between the two frames, the correlation coefficient r is almost 1 and we can neglect the second term in the equation. On the other hand, if there is significant motion, r will be small and we can no longer neglect the second term.

We use the variance of the frame difference to estimate the noise power σ_N^2. However, because the term $2\sigma_S^2(1-r)$ cannot be neglected when there is significant motion between two frames, we first filter the frame difference to remove all pixels in areas where there is motion. The filter used to do so is described next.

2.2. Structure detection filter

The filter we want to design must have the two following characteristics: it must be able to detect the pixels of moving areas, and it must not change the noise level of the sequence. To do so, we have developed a *structure detection filter.* This filter is based on the hypothesis that neighboring correlated pixels of the difference frame $D[x, y]$ represent a moving object and therefore cannot be used to estimate σ_N^2.

For each pixel, the filter looks at 4 neighboring groups of pixels, as depicted in Figure 2.

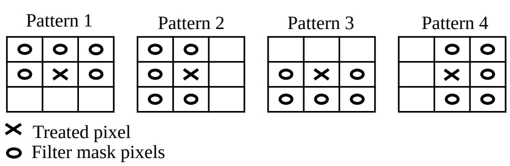


Figure 1. Patterns of the structure detection filter.

If the average of one of these groups differs significantly from the average of the difference frame $D[x, y]$, we assume that the pixel is part of a structure, and therefore that it is in a moving area. We assign to that pixel the mean value of $D[x, y]$ so that it will not be used in our evaluation of the noise variance.

We use β^2 to denote our estimator of σ_N^2, and its value is given by

$$\beta^2 = \frac{1}{2N}\sum_{x,y}\left(F(D[x,y]) - \hat{\mu}_d\right)^2 \tag{3}$$

where $F(D[x, y])$ is the output of the structure detection filter and is given by

$$F(D[x,y]) = \begin{cases} D[x,y] & \text{if } \forall i\text{: } \hat{\mu}_D - \alpha\hat{\sigma}_D < \rho_i < \hat{\mu}_D + \alpha\hat{\sigma}_D \\ \hat{\mu}_D & \text{otherwise} \end{cases} \tag{4}$$

where

$$\rho_n = \frac{1}{5}\sum_{[x,y] \in patern_n} D[x,y] , \tag{5}$$

N is the number of pixels in a frame, and $\hat{\mu}_D$ and $\hat{\sigma}_D$ are respectively the sample mean and sample variance of $D[x,y]$.

Note that the filter output at [x,y] does not depend on the pixel value at [x,y]. Since we assume that the additive noise $N[x, y]$ is white, the output of the detection filter does not change the noise power even if some pixels are wrongly assumed to be part of moving areas.

The results of the structure detection filter applied on the difference between two noisy frames is given in Figure 3. The constant α was set to 2 for all our tests. Note that the noise

(a)
(b)

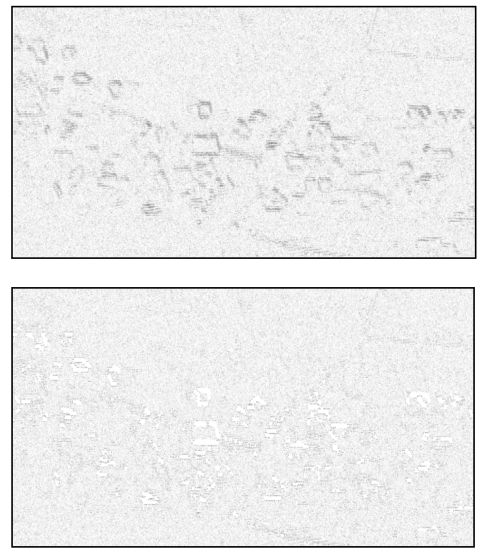


Figure 2. Frame difference of img. 74 and 75 of NATO sequence DIM10 (a) before and (b) after using the structure detection filter.

power estimate does not have to be computed for each frame since the noise level in a sequence typically is constant or varies slowly.

3. ADAPTIVE THRESHOLDING

3.1. Hierarchical block matching algorithm

As we mentioned earlier, the threshold we proposed is embedded in a hierarchical block matching algorithm (HBMA) which we will now briefly describe in order to introduce relevant notation. We use a simple two level HBMA, which means that both frames I_n and I_{n-1} are first low pass filtered and then subsampled. Then, for each block of $M/2 \times M/2$ pixels, a motion vector is computed using an exhaustive search block matching algorithm [2]. This is accomplished by minimizing with respect to (v_x, v_y), for a given block, the following expression:

$$\varepsilon'^2_{v_x v_y} = \frac{4}{M^2}\sum_{i,j}\Big(I'_{n-1}(x_0+v_x+i, y_0+v_y+j) - I'_n(x_o+i, y_o+j)\Big)^2 \tag{6}$$

where the symbol ' denotes low pass filtered and subsampled frames or quantities obtained with such frames. The exhaustive search is then pursued at full resolution, but this time for a smaller search area center around twice the vector value identified at the lower resolution.

3.2. Simple thresholding

Before estimating a motion vector, the following threshold test can be applied to determine if a block of pixels is stationary or if a significant motion vector can be estimated:

$$\varepsilon^2_{00} \le 2K\beta^2 \quad (7)$$

where K is a threshold constant and

$$\varepsilon^2_{00} = \frac{1}{M^2}\sum_{i,j} \left(I_{n-1}(x_0+i, y_0+j) - I_n(x_0+i, y_0+j) \right)^2 \quad (8)$$

We will later refer to the test of equation (7) (with K=1) as the zeroth order test. If the test is positive, the block is assumed to be similar to the one of the previous frame, and its motion vector is set to (0,0). If the test is negative, we then proceed with the estimation of the motion vector. This approach is similar to that in [5] except that the noise power is estimated from the data using equation (3).

Thus far, we have proposed a test but we have not discussed how to set the value of the constant K. We should first note that the mean squared difference, ε^2_{00}, using the image model described in the section 2.1, can be rewritten as

$$\varepsilon^2_{00} = 2\hat{\sigma}^2_N + 2\hat{\sigma}^2_S(1-\hat{r}) + (\hat{\mu}_{I_n} - \hat{\mu}_{I_{n-1}})^2 \quad (9)$$

where the symbol ^ is used here to denote sampled quantities computed over a block of pixels as opposed to statistical expectation.The first term of this equation is the noise contribution, where the last two are the components due to block motion. It is the relative importance of the first term with respect to the last two that should determine the value of the constant K

If the dominant terms are the last two, the threshold constant should be small to take advantage of the motion compensation. If, on the contrary, the dominant term is $2\hat{\sigma}^2_N$, we would like the constant K to be large since motion compensation is unlikely to generate a better prediction. In noisy sequences, this may lead to a significant reduction of the computation time since motion estimation will be performed for fewer vectors.

Taking advantage of the different spectral characteristics of the terms in (9), we will now present a simple test that will have the behavior we have just described.

3.3. Implicitly adaptive threshold test

We propose the following test which implicitly adjusts the threshold constant of equation (7):

$$\varepsilon'^2_{00} \le 2\beta^2 \quad (10)$$

where ε'^2_{00} is a block mean squared difference computed from low pass filtered and subsampled frames. We will refer to this test as the first order test since it is based on subsampled images. Because the noise contribution to (9) is typically white, while the motion contribution is correlated (*i.e.*, has a low pass spectrum), the low pass filtering of the frame difference will reduce significantly the energy due to the noise compared to the energy due to pixel displacement.

If there is no motion in the block, we can show that for white noise with a first order low pass Gaussian filter

$$\varepsilon'^2_{00} = 0.141\varepsilon^2_{00}. \quad (11)$$

This means that the test of equation (10) is equivalent to the following test

$$\varepsilon^2_{00} \le \frac{2\beta^2}{0.141}. \quad (12)$$

On the other hand, if the motion within a block is significant, the frequency content of the frame difference will be in the lower bands, as compared to the noise spectral distribution. This means that the frame difference energy before and after filtering will be similar. For the special case of an object of constant value moving rapidly over a background of another constant value, the frame difference will be constant, which means that $\varepsilon^2_{00} = \varepsilon'^2_{00}$. In this case, the test of equation (10) is

$$\varepsilon^2_{00} \le 2\beta^2 \ . \quad (13)$$

The two limit cases we have just described show that the first order test is equivalent to the test of equation (7) where the constant K is adaptively changed to reflect the importance of the frame difference energy due to motion compared to that due to the noise.

When the noise energy is high compared to the energy due to motion, the energy of the total prediction error must be significantly smaller than the noise energy if we want to observe a coding gain by using a nonzero motion vector. That is exactly what the first order test is doing by reducing the value ε'^2_{00} to a smaller value than that of ε^2_{00}. On the other hand, when the noise energy is low compared to that due to the motion, the value of ε'^2_{00} is kept close to that of ε^2_{00}, which means that the outcome of the first order test is more likely to be negative and motion compensation of the block will take place.

4. RESULTS

4.1. In noisy sequences

We have applied the proposed threshold test on both infrared and standard video sequences. The infrared sequences are part of a NATO set showing dim targets moving slowly. These sequences are characterized by a high level of noise. Figure 3a shows a frame of one of these sequences. In this example, the sequence shows a village in the foreground with two dim targets in the background. The targets are moving, and the camera is panning the scene.

The motion field obtained with a standard BMA is shown in Figure 3b. Figures 3c and 3d respectively show the motion field obtained when the zeroth and first order threshold tests are used prior to motion estimation. These results show clearly that the motion field obtained with the first order threshold test is closer to the scene's true motion. Most of the nonsignificant vectors have been zeroed out while the motion vectors in the target areas have been correctly estimated. This was not the case with the zeroth order test for which a significant number of nonsignificant vectors have been estimated.

4.2. In noiseless sequences

We present in Figure 4 results obtained on the video sequence *football*. Figures 4a and 4b respectively show a frame of the sequence, and the motion field obtained with a standard BMA. The motion fields obtained with both threshold tests are shown in Figures 4c and 4d.

Note that in this case, the motion fields obtained with both threshold tests are similar. This was expected since for this sequence, the noise level is small and the frame difference energy is dominated by the motion of the players. These tests therefore show that, as we argued, the energy of the frame difference before and after filtering does not differ significantly.

The motion field obtained with the threshold tests differ from the one computed by BMA in the textured areas. The structure detection filter used to eliminate moving areas before estimating the noise power cannot identify the textured areas as moving. These textured areas are therefore used to estimate the noise power, and then are treated as such by the threshold tests. This behavior should not cause any problems since the human visual system is not sensitive to texture. Coding errors in these areas can therefore be more severely quantized. Further tests have to be done to assess more precisely this effect.

However, as presented, the proposed test can also be used to initialize a segmentation of moving objects from a textured background. This type of segmentation is typically part of object-oriented coders like the one proposed in MPEG4.

5. CONCLUSION

We have presented in this paper an adaptive motion detection threshold test. This test takes advantage of the difference between the spectral content of noise and the spectral content of structures generated by motion in a frame difference to implicitly adjust a threshold constant. The proposed test, embedded in a hierarchical block matching algorithm, can significantly speedup the motion estimation by assigning the zero vector to blocks dominated by noise and for which no significant motion vectors can be estimated. As shown by the results, this algorithm is particularly efficient in noisy sequences. As presented, this algorithm can also be used with standard video sequences, typically characterized by a lower noise level. In this case, the algorithm can be used to initialize a segmentation of moving objects from textured backgrounds.

By replacing the noise power estimate by that of the energy of the quantization noise after the encoding of the displaced frame difference, the proposed algorithm can also be used in an MPEG coder to adaptively adjust the threshold test used to decide on the coding mode of a macroblock.

6. REFERENCES

[1] E. Salari and Sheng Lin, "Segmentation based coding algorithm for low bit-rate video", *Proceeding of IEEE ICASSP*, pp. 2241-2244, 1995.

[2] J.R. Rain, "Displacement measurement and its application in interframe image coding", *IEEE Trans. Commun.*, vol-29, no. 12, pp. 1799-1808, 1981.

[3] M. Bierling, "Displacement estimation by hierarchical block matching". *SPIE Visual Commun. and Image Proc.*, vol. 1001, pp. 387-403, 1988.

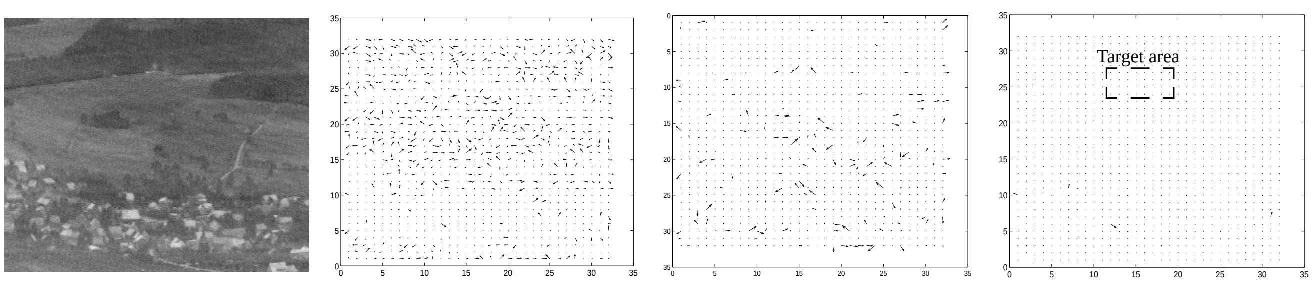
[4] N. Dielh, "Object-oriented motion estimation and segmentation in image sequences coding", *Signal Processing: Image Communication,* vol. 3, no. 1, pp. 23-56, February 1991.

[5] V. Bhaskaran, W. Li and M. Kunt, "A segmentation based scheme for very low bit-rate video coding", *SPIE Digital Video Compression*, vol. 2419, pp. 81-89, 1995

[6] MPEG D.I.S. ISO/IEC 11172, 1992.

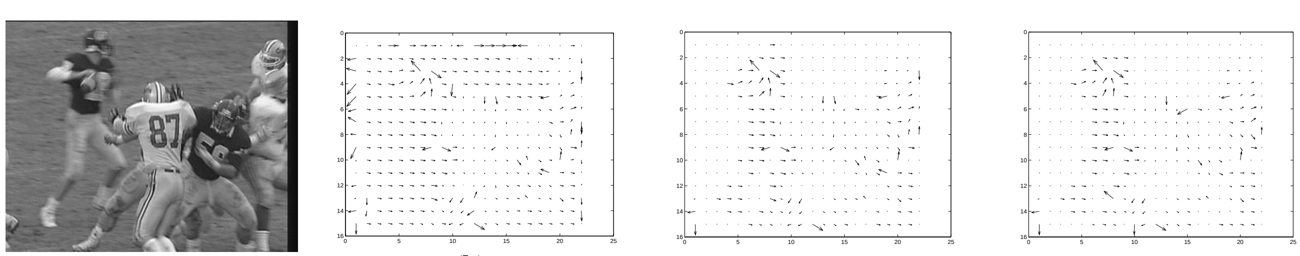
7. ACKNOWLEDGMENT

This work was supported by a contract from the Canadian Defense Research Establishment at Valcartier.



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

Figure 3. Motion field between images 74 and 75 of the NATO sequence DIM10. (a) Image 74, (b) BMA, (c) Zeroth order test (k=1), (d) First order test.



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

Figure 4. Motion field between football images 135 and 136. (a) Image 135, (b) BMA, (c) Zeroth order test (K=1), (d) First order test.