

HMM BASED METHOD FOR DYNAMIC TEXTURE DETECTION

B. Ugur Toreyin, A. Enis Cetin
Bilkent University 06800 Ankara Turkey
{ugur,cetin}@ee.bilkent.edu.tr

ABSTRACT

A method for detection of dynamic textures in video is proposed. It is observed that the motion vectors of most of the dynamic textures (e.g. sea waves, swaying tree leaves and branches in the wind, etc.) exhibit random motion. On the other hand, regular motion of ordinary video objects has well-defined directions. In this paper, motion vectors of moving objects are estimated and tracked based on a minimum distance based metric. The direction of the motion vectors are then quantized to define two three-state Markov models corresponding to dynamic textures and ordinary moving objects with consistent directions. Hidden Markov Models (HMMs) are used to classify the moving objects in the final step of the algorithm.

1. INTRODUCTION

Two dimensional (2-D) textures and related problems were extensively studied in the field of computer vision [Forsyth(1)]. On the other hand, there is very little research on three-dimensional (3-D) texture detection in video [2, 3]. Trees, fire, smoke, fog, sea waves, sky and shadows are examples of time-varying 3-D textures in video. It is well known that tree leaves in the wind, moving clouds, etc., cause major problems in outdoor video motion detection systems. If one can initially identify bushes, trees and clouds in a video, then such regions can be excluded from the search space or proper care can be taken in such regions. This leads to robust moving object detection and identification systems in outdoor video. In this paper, a method for detection of tree branches and leaves in video is proposed.

Motion detection in video constitutes the primary step for almost all types of video based surveillance applications. [4]. It is observed that, the motion vectors of tree branches and leaves exhibit random motion. On the other hand, regular motion of green colored objects has well-defined directions. In this paper, the wavelet transform of motion vectors are computed and objects are classified according to the wavelet coefficients of motion vectors. Color information is also used to reduce the search space in a given image frame of the video. Motion

trajectories of moving objects are modeled as Markovian processes. In the final step of the algorithm, Hidden Markov Models (HMMs) are used to classify the green colored objects according to their motion trajectories.

In Section 2, detection algorithm is described. In Section 3 experimental results are presented.

2. DETECTION ALGORITHM

Our detection algorithm consists of three main steps: i) green colored moving region detection in video, ii) analysis of the motion trajectories in the wavelet domain, and iii) HMM based classification of the motion trajectories.

2.1. Moving region detection

Moving pixels and regions in the video are determined by using a background estimation method developed in [5] in which camera monitoring the scene is assumed to be stationary. In this method, a background image B_{n+1} at time instant $n+1$ is recursively estimated from the image frame I_n and the background image B_n of the video as follows:

$$B_{n+1}(k,l) = \begin{cases} aB_n(k,l) + (1-a)I_n(k,l), & \text{if } I_n(k,l) \text{ stationary} \\ B_n(k,l), & \text{if } I_n(k,l) \text{ moving} \end{cases} \quad (1)$$

where $I_n(k, l)$ represents a pixel in the n^{th} video frame I_n , and a is a parameter between 0 and 1. Moving pixels are determined by subtracting the current image from the background image and adaptive thresholding (cf. Fig. 1a). For each pixel an adaptive threshold is estimated recursively in [5]. Pixels exceeding thresholds form moving regions and they are determined by connected component analysis.

We do not need very accurate boundaries of moving regions. Hence the above computationally efficient algorithm is sufficient for our purpose of estimating the motion vectors of green colored moving regions in video. Other methods including the ones described in [6] and [7]

can also be used for moving pixel estimation but they are computationally more expensive than [5].

We are solely concentrated on the detection of swaying leaves in video, therefore we incur a simple color constraint, $G > B$, on green (G) and blue (B) channels of the RGB color space to reduce the size of the search space.



Fig. 1. (a) Moving pixels, and (b) their minimum bounding boxes are determined

2.2. Analysis of motion trajectories in wavelet domain

After a post-processing stage comprising of connecting the pixels, moving regions are encapsulated with their minimum bounding rectangles (cf. Fig.1b). Next, these moving regions in the current frame are matched to the closest moving regions in the previous frame. Euclidean metric is used for distance calculation. A motion trajectory is kept for each moving region.

Tree branches and leaves usually exhibit a swaying motion trajectory which has a dominant horizontal (x) component compared to its vertical (y) component. The magnitude of these vectors are smaller than the motion vectors of regular moving objects. Another difference between the motion characteristics of swaying leaves and regular green colored moving objects is that regular moving objects have well-defined directions throughout the course of their motion. However, tree leaves in the wind sways back and forth within a limited region without a sense of particular direction (cf. Fig.2).

Therefore, we only make use of temporal variations in the x -component of motion vectors and analyze them in the wavelet domain. For each moving region, n frame horizontal motion vector history is kept for its center of mass. The vectors in the motion history are quantized according to their directions as we are only concerned with the directions of the vectors. Hence we have the quantized motion feature signal $u_x = -1, 0, 1$. The values +1 and -1 correspond to opposite directions in the horizontal axis, and the zero value means the object under consideration is stationary or below a threshold.

The temporal variations in the x -components of the center of masses of the leaves and the car in Fig.2, are presented in Fig.3(a) and Fig.4(a), respectively. Defining

the horizontal direction from right to left as positive direction, the temporal variation in the quantized motion feature signal of the car and the leaves are shown in Fig.3(b), and Fig.4(b), respectively.



Fig. 2. The car has a directionally consistent trajectory whereas the leaves, pointed with an arrow, sway randomly in the wind

We then calculate the corresponding wavelet coefficients for this ternary motion feature signal, u_x . Wavelet coefficients, w 's, are obtained by high-pass filtering followed by decimation as shown in Fig.5.

The wavelet transform of the one-dimensional motion signal is used as a feature signal in HMM based classification in this paper. It is experimentally observed that this feature signal exhibits different behavior for the leaves swaying in the wind and the objects with directionally consistent trajectories. A random behavior with low temporal correlation is apparent for leaves and branches of a tree, in both the horizontal component of the temporal motion signal and its corresponding wavelet signal as shown in Figs. 3(c) and 4(c), respectively. On the other hand, an ordinary green moving object with a well-defined direction does not exhibit such a random behavior. In this case there is high correlation between the samples of the motion feature signal. This difference in motion characteristics is also apparent in the wavelet domain.

The use of wavelet coefficients, w 's, instead of the quantized motion vector to characterize moving regions has some major advantages over the use of actual temporal signal. The primary advantage is that, wavelet signals can easily reveal the highly correlated characteristic of the motion which is intrinsic for directionally consistent moving objects. The unit direction vectors for these objects are the same for all the time during their entire travel, except for some turns with

horizontal component of their motion vector reversed. Since, wavelet signals are high-pass filtered signals, no variations in the original signal lead to zero wavelet coefficients. Hence it is easier to set thresholds in the wavelet domain which are robust to variations of trajectories. Wavelet coefficients of the motion signal of tree branches is also a zero mean signal but its variance is high due to high-frequency nature of the original temporal motion signal.

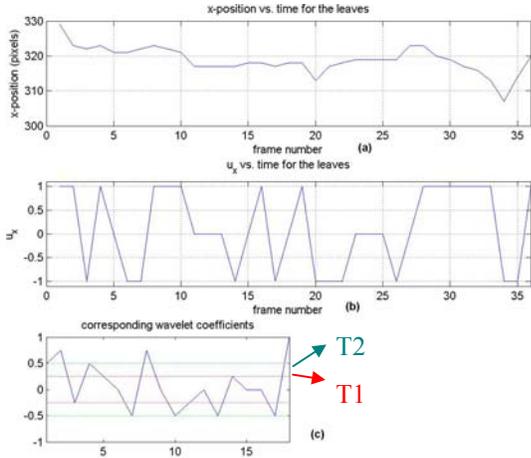


Fig. 3. (a) x-position variation with time of the center of mass of the leaves in Fig.2, (b) corresponding quantized motion signal, and (c) the wavelet coefficients of the quantized motion signal.

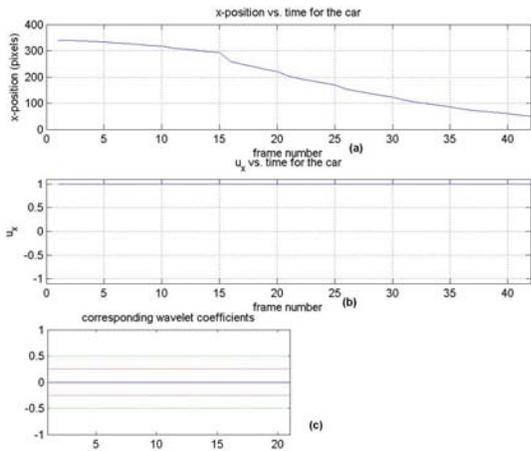


Fig. 4. (a) x-position variation with time of the center of mass of the car in Fig.2, (b) corresponding quantized motion signal, and (c) its wavelet coefficients. Since the car in Fig.2 does not change its direction in the horizontal axis, there is no variation in signals shown in (b) and (c)

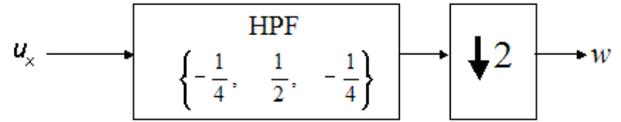


Fig. 5. Wavelet coefficients, w corresponding to motion feature signal, u_x , are evaluated with an integer arithmetic high-pass filter (HPF) corresponding to Lagrange wavelets [8] followed by decimation

We set two thresholds, T1 and T2 for defining Markov states in the wavelet domain as shown in Fig. 3. The lower threshold T1 basically determines the wavelet signal being close to zero. For a direction consistent moving object, normally the motion vector is constant except for a few number of turns. When the motion direction is consistent, the wavelet signal is zero. After a few turns, the wavelet signal start taking values close to zero. When turning is over, it takes zero values again. The use of wavelet domain information also makes the method robust to subsequent variations in the direction of the moving object's trajectory. This is achieved by the use of the second threshold T2 to detect high amplitude variations in the wavelet signal, which correspond to edges or high-frequency changes in the original signal. When the wavelet coefficients exceed the higher threshold T2 in a frequent manner this means that the object is changing its direction or exhibiting periodic behavior due to bending or swaying back and forth in the wind.

2.3. HMM based classification

Regular motion of the green colored objects exhibits a Markovian behavior with strong correlation. On the other hand, horizontal component of the motion vector of tree branches have little correlation in time. Therefore, Markov model based classification is ideal for the classification problem.

Two three-state Markov models are used to classify the motion of objects in this paper. Non-negative thresholds $T1 < T2$ introduced in wavelet domain, define the three states of the Hidden Markov Models for leaves and directionally consistent moving objects as shown in Fig.6(a) and (b), respectively.

At time n , if $|w(n)| < T1$, the state is in $S1$; if $T1 < |w(n)| < T2$, the state is $S2$; else if $|w(n)| > T2$, the state $S3$ is attained. During the training phase of the HMMs transition probabilities a_{uv} and b_{uv} , $u, v = 1, 2, 3$, for leaves and directionally consistent moving object models are estimated off-line, from a set of training videos. In our experiments, 20 consecutive image frames are used for training HMMs.

For the leaves, since the motion is quasi-periodic, we expect similar transition probabilities between the states. Therefore the values of a 's are close to each other.

However, for directionally consistent moving objects, the wavelet signal takes values around zero. Hence we expect a higher probability value for b_{00} than any other b value in the directionally consistent moving object model, which corresponds to higher probability of being in S1. The state S2 provides hysteresis and it prevents sudden transitions from S1 to S3 or vice versa.

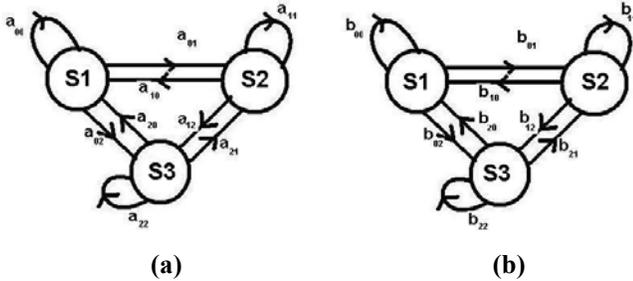


Fig. 6. Three state Markov models for (a) leaves, and (b) directionally consistent moving objects

During the recognition phase the state history of length 20 image frames are determined for the moving objects detected in the viewing range of the camera. This state sequence is fed to the leaves and directionally consistent moving object models. The objects for which the directionally consistent moving object model yields higher probability are suppressed. Only the moving objects for which leaves model yield higher probability is kept. The pixels for which color constraint is satisfied within these moving objects form the leaves mask.

3. EXPERIMENTAL RESULTS

The proposed algorithm works in real-time on an AMD AthlonXP 2000+ 1.66GHz processor. As described above HMMs are trained from outdoor video clips with swaying tree leaves in the wind and regular moving objects. A total of 12 video clips having 5633 image frames with 360x280 pixel resolution are used. Four of the clips are captured at 5 fps and the others have capture frame rate of 10 fps.

We trained our models with two of the clips having both tree leaves in the wind and regular moving objects, such as cars and walking people. The remaining ten clips are used for test purposes. Our method yields no false positives in any of the clips.

Detection results for test videos are presented in Table1. There are parking cars and walking people in almost all of the test video clips. Image frames from some of the clips are shown in Fig.7. Our method detects leaves that are persistently swaying in the wind for a while. It does not detect leaves that move in a few frames. This is mainly due to the fact that we need to build a Markovian model of the motion and this obviously requires a temporal history of the motion. Once tree branches and leaves are identified, their locations in the video are

determined by the surveillance system and random motion in such regions can be excluded to eliminate false alarms due to the motion of tree branches in the wind.

Table1: Detection results for ten test videos. The middle column lists the number of frames in which there is motion due to moving tree leaves in the wind. The column on the right shows the number of frames in which tree leaves are detected by our method

CLIPS	Number of frames in which leaves sway in the wind	Number of frames in which leaves detected with our method
V1	0	0
V2	0	0
V3	70	47
V4	45	36
V5	35	13
V6	9	2
V7	0	0
V8	74	42
V9	617	502
V10	107	43

4. CONCLUSION

A method for detection of swaying tree branches and leaves in video is proposed. Random motion of tree branches and leaves in the wind are used to recognize the tree branches and leaves. The wavelet transform of motion vectors are computed and objects are classified according to the wavelet coefficients of motion vectors. Regular motion of ordinary green colored objects exhibits a Markovian behavior with strong correlation. On the other hand, horizontal of the motion vector of tree branches have little correlation in time. The use of wavelet coefficients instead of actual motion vectors in an HMM framework for classification provides more robust results.

5. REFERENCES

[1] Forsyth, D.A., and J. Ponce *Computer Vision – A Modern Approach*, Prentice Hall, 2002.
 [2] Y. Dedeoglu, B.U. Toreyin, U. Gudukbay, and A.E. Cetin, "Computer Vision Based Method for Real-time Fire and Flame Detection," in *Proceedings of IEEE ICASSP'05*, p.669-673, 2005.

[3] W. Phillips III, M. Shah, and N.V. Lobo, "Flame Recognition in Video," *Pattern Recognition Letters*, Elsevier, vol. 23 (1-3), pp. 319-327, 2002

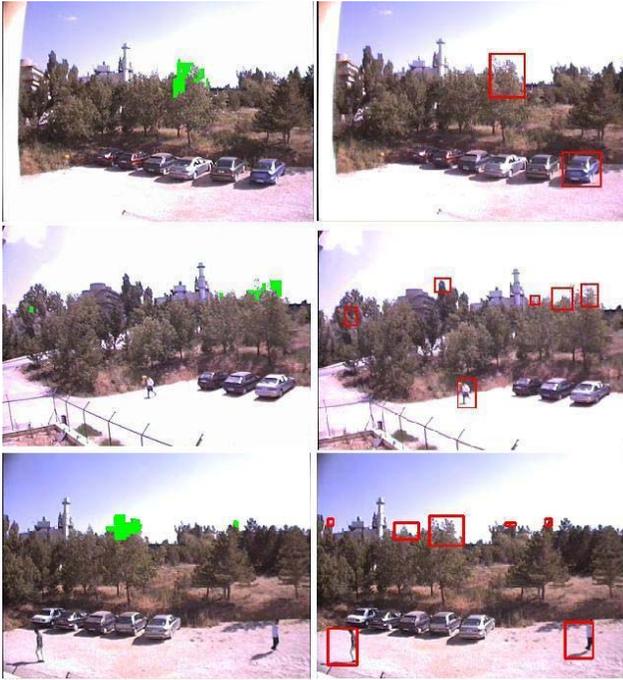


Fig. 7. Sample image frames from some of the test video clips. The images on the left are the detection results of our method. Detected leaves are in green. The images on the right show all the moving objects present in the scene.

[4] C. Regazzoni, V. Ramesh, G.L. Foresti, "Scanning the Issue/Technology," *Proceedings of the IEEE*, vol. 89 (10), pp. 1355-1365, 2002.

[5] R.T. Collins, A.J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L. Wixson, "A System for Video Surveillance and Monitoring: VSAM Final Report," *Tech. report CMU-RI-TR-00-12*, Carnegie Mellon University, 2000.

[6] M. Bagci, Y. Yardimci, A.E. Cetin, "Moving Object Detection Using Adaptive Subband Decomposition and Fractional Lower Order Statistics in Video Sequences," *Signal Processing*, Elsevier, p.1941-1947, 2002.

[7] C. Stauffer, W.E.L. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 246-252, 1999.

[8] C.W. Kim, R. Ansari, A.E. Cetin, "A class of linear-phase regular biorthogonal wavelets," in *Proceedings of IEEE ICASSP'92*, p.673-676, 1992.