2010 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2010) December 6-8, 2010

LOW RESOLUTION METHOD USING ADAPTIVE LMS SCHEME FOR MOVING OBJECTS DETECTION AND TRACKING

Chih-Hsien Hsia, Yi-Ping Yeh, Tsung-Cheng Wu, Jen-Shiun Chiang, and Yun-Jung Liou

Department of Electrical Engineering, Tamkang University, Taiwan
E-mail: chhsia@ee.tku.edu.tw; rlpolo123@hotmail.com; chiang@ee.tku.edu.tw

ABSTRACT

This paper presents a new model for adaptive filter with the least-mean-square (LMS) scheme to train the mask operation on low resolution images. The adaptive filter theory with adaptive least-mean-square scheme (ALMSS) uses the training mask for moving object detection and tracking. However, the successful moving objects detection in a real surrounding environment is a difficult task due to noise issues such as fake motion or Gaussian noise. Many approaches have been developed in constrained environments to detect and track moving objects. On the other hand, the ALMSS approach can effectively reduce the noise with low computing cost in both fake motion and Gaussian noise environments. The experiments on real scenes indicate that the proposed ALMSS method is effective for moving object detection and tracking in real-time.

1. INTRODUCTION

Autonomous video surveillance and monitoring have been developed for long time, and various approaches to track moving objects have been proposed in recent years [1]-[2]. To correctly track the moving object does not only provide an accuracy results for post-processing, but also can reduce the redundant computation for the incorrect detection of the moving object. However, moving object detection in a real surrounding environment is essential and needs to be further improved owing to some difficult situations, such as illumination changes, fake motions, night detection, and Gaussian noise in the background. All these situations may lead to detect incorrect region of the moving object. For solving these problems, several methodologies have been proposed such as background subtraction, temporal differencing, and optical flow for object detection [1]-[2].

Cheng *et al.* [3] used discrete wavelet transform (DWT) to detect and track moving objects. The 2-D DWT can be used to decompose an image into four-subband images (LL, LH, HL, and HH). The LL-band image produced by the original image size via two dimensions (row and column) calculation may cause high computing cost in the pre-processing. After dealing with the background subtraction, Alsaqre *et al.* [7] used a local pre-process method to smooth the image with reducing noise and other small fluctuations. However, this approach is unable to reduce the post-processing computation. Sugandi *et al.* [4] proposed a method for detecting and tracking objects by using a low resolution image with the 2×2 average filter. They mentioned that the low resolution image is insensitive to illumination changes and can reduce the small movement like moving leaves of trees in the background. Although this method can deal with small movement, these low resolution images become more blurred than the LL-band image generated by using DWT.

To overcome the above-mentioned problems, we propose a method based on adaptive filter theory, adaptive least-mean-square based scheme (ALMSS), for detecting and tracking moving objects. In ALMSS, we can select only the LL-band model. Unlike the conventional DWT method to process row and column dimensions separately by low-pass filter and down-sampling, the LL-band of ALMSS can be used to directly calculate the LL-band image. Our method is more robust because it can change the coefficients of the mask in accordance with the dynamic environment. The method can also reduce the image transfer computing cost and remove fake motions not belonged to the real moving objects.

The rest of this paper is organized as follows. In Section Ⅱ, we briefly introduce the general flow chart for detecting and tracking moving objects by using low resolution image. The proposed method and the efficient model for moving object tracking are presented in Section Ⅲ. Section Ⅳ shows the experimental results. The conclusions are given in Section Ⅴ.

2. MOVING OBJECTS TRACKING USING LOW RESOLUION METHOD

Images are often corrupted by noises. The noises may be produced by not only the real environmental noise but also the imperfection of video acquisition systems and transmission channels. Consequently, this will cause a significant reduction of video quality, especially for the high-level computer vision. Before dealing with motion objects tracking, there are certain methods for removing noise or fake motion with low computing cost proposed in the past several years [3], such as the low resolution method.

978-1-4244-7371-7/10/$26.00 ©2010 IEEE

2.1. Low Resolution Image

Sugandi *et al.* [4] proposed a simple method by using the low resolution concept to cope with the noise such as fake motion or small movement in the background. They generate the low resolution image by replacing the pixel value of an original image with the average value of its four neighbor pixels and itself. The average filter is a low pass filter which can reduce the noise and perform restoration by the noise reduction in the spatial domain. It also provides a flexible multi-resolution image like that of DWT.

2.2. Detection and Tracking Flow

The pre-processing flowchart of the low resolution image for moving object detection and tracking system is shown in Fig. 1. Basically we apply the frame difference method to detect the moving objects. In order to decrease the holes left inside the moving entities, three continuous frames (F_{t-1}, F_t, and F_{t+1}) are used in this system for detecting moving object mask. These three continuous frames are decomposed into LL_2-band frames (LL_{2t-1}, LL_{2t}, and LL_{2t+1}) by using a low resolution image. It can proceed with the post-processing by employing these three LL_2-band frames. Binary masks, B_{t-1} and B_t, can be obtained by computing the binary values of these three successive LL_2-band frames (in between LL_{2t-1}, LL_{2t}, and LL_{2t+1}) and a threshold value T in (1).

$$B_{t-1}(i,j) = \begin{cases} 1, & \text{if } |LL_{2t-1}(i,j) - LL_{2t}(i,j)| > T \\ 0, & \text{otherwise} \end{cases}$$

$$B_t(i,j) = \begin{cases} 1, & \text{if } |LL_{2t}(i,j) - LL_{2t+1}(i,j)| > T \\ 0, & \text{otherwise} \end{cases} \qquad (1)$$

The motion mask (MM_t) can be generated by using the union operation (logical OR) of B_{t-1} and B_t. The operating function is represented as follows:

$$MM_t = B_t \cup B_{t-1}. \qquad (2)$$

The holes may still exist in the motion masks, because some motion pixels are too tiny such that it causes error judgments as non-motion ones. In order to increase the motion mask (MM_t) robustly, the morphological *closing* method is used to fill these holes. First, we apply the dilation operator for filling the middle of the isolated pixels that become related in the motion masks. It is defined as follows:

$$F_t(i,j) = \begin{cases} 1, & \text{if one or more pixels of the adjacent pixels of motion mask } MM_t(i,j) \text{ are } 1, \\ 0, & \text{otherwise.} \end{cases} \qquad (3)$$

Then we apply the erosion operator for eliminating redundant pixels in the motion mask boundary as follows:

$$MMR_t(i,j) = \begin{cases} 0, & \text{if one or more pixels of the adjacent pixels of motion mask } F_t(i,j) \text{ are } 0, \\ 1, & \text{otherwise.} \end{cases} \qquad (4)$$

It scans eight neighbors of the motion mask MMRt image pixel by pixel from top left to bottom right. After extracting the connected component, it obtains several moving objects. In this work, we utilize the region-based tracking algorithm [3] to track the moving object motion.

Labeling is useful when the moving objects in the scene are more than one (The connected component labeling is then employed to label each moving object and track each moving object individually). The labeled moving objects are thus found, and then we extract the boundary of the moving object using rectangle box to track the moving object. In order to track moving objects in the original image size, we have to transform the coordinate from the LL_2 image size back to the original image size according to the spatial relationship of the LMS algorithm which depends on DWT as follows:

$$O(x,y) = LL_n(x \times 2^n, y \times 2^n). \qquad (5)$$

where n is the order of LL-band frames.

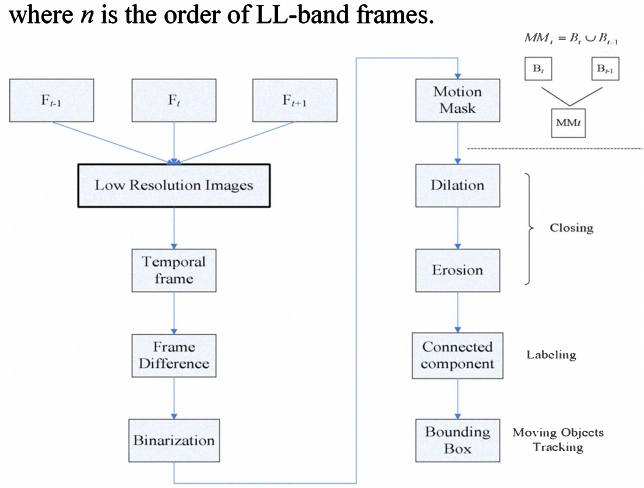


Fig. 1. Flow chart of the moving object detection and tracking by using the low resolution image.

3. ADAPTIVE LEAST MEAN SQUARE BASED SCHEME

In order to detect and track the moving object more accurately, we propose a new method based on adaptive filter theory with adaptive least-mean-square scheme (ALMSS). It preserves more image quality of the low resolution image than that of the low resolution method [4] to reduce the spatial resolution of the image by replacing the general mask operation, such as average filter, down-sampling, and DWT, with the LMS-based mask.

3.1. Least Mean Square Algorithm

Adaptive filters constitute an important part of the statistical signal processing. There are many iterative search algorithms derived to find the optimum tap-weights (the coefficients of mask operation), such as minimizing the mean-square-error (MSE). In this work, we use LMS for automatically adjusting the equalizer coefficients to optimize a specified performance index. An N-tap transversal adaptive filter can be described in Fig. 2.

In Fig. 2, $x(n)$ is the signal input; $y(n)$ is the signal output, and we need the desired output $d(n)$ to consider the error estimation $e(n)$, which minimizes the MSE for finding the optimum tap-weights vector $W(n)$ as the following derivations:

$$y(n) = \sum_{i=0}^{N-1} w_i(n)x(n-i) \quad (6)$$

$$e(n) = d(n) - y(n) \quad (7)$$

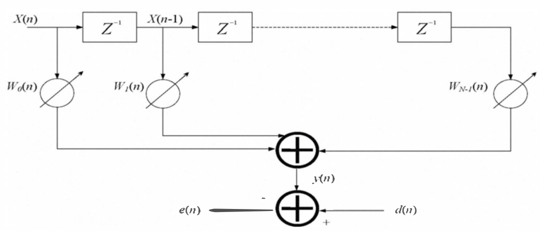


Fig. 2. N-tap transversal adaptive filter.

Finally, we obtain (8) as the LMS recursive operation:

$$\bar{W}(n+1) = \bar{W}(n) + 2\mu e(n)\bar{x}(n) \quad (8)$$

where $\mu > 0$ is called the step-size parameter, which is selected by the autocorrelation matrix of the filter inputs. In other words, the tap-weights can converge to an optimum result if and only if the step-size parameter μ is selected as

$$0 < \mu < \frac{1}{\lambda_{\max}} \quad (9)$$

In (9), $\lambda_{\max}$ is the maximum eigenvalue of the autocorrelation matrix $E[\bar{x}(n)\bar{x}^T(n)]$, which has a relationship of the input signal $x(n)$. The LMS algorithm adapts the filter tap weights for the effect minimized in the MSE [6].

3.2. Least Mean Square Algorithm for Moving Objects Tracking

The proposed method for moving objects tracking uses a low resolution image with a 5×5 symmetric mask-based filter based on LMS. To obtain the coefficients of the mask, we build the training flow by using LMS with DWT operation as shown Fig. 3. The LMS method is an iterative gradient-descent based procedure. In this case, the method can converge the result to a unique global solution.

The training process involves simultaneously presenting the input vector and desired output to the adaptive linear transversal filter and modifying the tap-weight vector until the output is as close as possible to the desired output. Therefore, under the LMS-based method, the process has to compare the output data with the desired data. The LMS approach consists of four steps: Step 1 obtains the desired vector (from the left hand side); Steps 2-3 are for the output vector (at the right hand side), and the compared results are generated in Step 4.

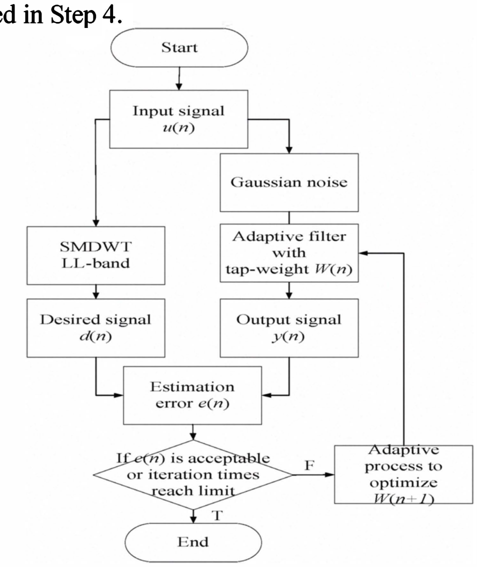


Fig. 3. The training flow by using LMS with DWT operation.

Step 1: Preparing a sequence of the input images which we are interested in. In order to presume the desired images, the SMDWT [5] of low resolution images for desired data is used.

Step 2: Joining the white Gaussian noise with the original input images which are polluted.

Step 3: Using the noisy images generated from Step 2 and calculating the output vector by using the LMS approach. We use the LL_2-band mask with 5×5 window size of the tap-weights.

Step 4: Correcting the weights in the output vector and the desired vector, where the output vector is comprised by the LMS output data and DWT output data as the desired vector shown in Fig. 4. Go back to Step 3 and repeat the calculation and correction until the objective function reaches the stop standard or the largest training times.

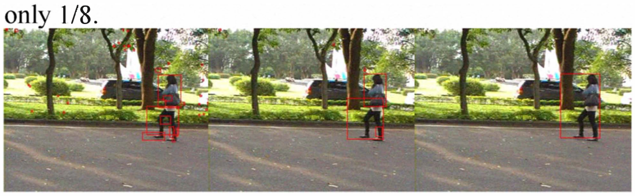
When the training flow is done, we may obtain a 5×5 symmetric mask-based filter. The original frames can be departed into low resolution frames by the mask operation. Finishing the above procedure, we can obtain a sequence of data for the motion detection by replacing the low resolution method as shown in Fig. 1.

4. EXPERIMENTAL RESULTS

In this work, the experimental results of the outdoor (all day) environment with statistic video system are demonstrated. The frame size of the video sequence is 320×240, and the format of color image frame is 24-bit in a RGB system. We use all gray level frames from transferring the RGB system to the YCbCr system for detecting moving object motion and utilize the LL_2 image size of 80×60 generated by using the LMS approach from the original image for our proposed moving object detection and tracking system. The experimental environment is set with 2.4 GHz Intel Core 2 Duo CPU, 2 GB RAM, Microsoft Windows XP SP3, and Borland C^{++} Builder 6.0. The software includes verifying for algorithms and image process for the moving objects detection.

4.1. The Low Resolution Results

The LMS method is applied to find low resolution results by using LL_1- and LL_2-band images, respectively, and the low resolution results are shown in Fig. 4. When dealing with several noises such as moving leaves of trees, the original image and the LL_1-band image have poor results because these noises sometimes are large that cannot be eliminated completely. Our method can successfully track the moving objects with reducing noise when applying the LL_2-band image. Our method can keep detecting and tracking the moving objects even if the resolution factor is only 1/8.



(a) (b) (c)

Fig. 4. Results of tracking moving objects in various time with static background: (a) original frames without LMS-based moving objects tracking method, (b) LL_1-band frames with LMS-based moving objects tracking method, (c) LL_2-band frames with LMS-based moving objects tracking method.

4.2.Moving Object Tracking

Without the LMS scheme, many noise regions are tracked. However, even if the moving objects are tracked, those moving regions are fragmented, as shown in Fig. 4(a). By using the LMS mask scheme, the noises can be filtered out, as shown in Fig. 4(b).

It still generates incomplete moving object regions by using the LL_1-band image, because the relevance of these pixels in the LL_1-band image is deleted. Finally, let us look at the results of the LL_2-band image in Fig. 4(c). Using the two-level band image has a better tracking region and also can cope with noises and fake motions effectively, as shown in Table I .

TABLE I. THE MOVING OBJECTS DETECTION AND TRACKING RESULTS

Resolution	[1]ALMSS	[2]LS	[3]2×2AFS	[4]DSS
Level	DTF/AR	DTF/AR	DTF/AR	DTF/AR
LL2 (80×60)	21 FPS/ 81.45 %	12 FPS/ 77.02 %	33 FPS/ 68.15 %	32 FPS/ 15.32 %

[1]LMS: LMS-based Mask Scheme; [2]LS: Lifting Scheme; [3]2×2 AFS: 2×2 Average Filter Scheme; [4]DSS: Down-Sampled Scheme; DTF: Detection + Tracking flow; AR: Accuracy rate.

5. CONCLUSIONS

The ALMSS based on adaptive filter theory for moving object detection and tracking is proposed in this paper. It is able to detect and track moving objects in complexity environments with statistic video systems. The proposed ALMSS does not only overcome the drawbacks of high complex computation and slow speed for the conventional DWT, but also preserves the wavelet features of the flexible multi-resolution image and the capability for dealing with noises and fake motions such as moving leaves of trees. The experimental results demonstrate that the LL_2-band (image size of 80×60) can effectively track moving objects by region-based tracking under any environments, as well as it can cope with noise issues. The ALMSS approach can be extended to the real-time video surveillance system applications, such as object classification and descriptive behaviors of objects.

6. REFERENCES

[1] W.-M. Hu, T.-N. Tan, L. W, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE TSMC- Part C: Applications and Reviews*, vol. 34, no. 3, pp. 334-352, Aug. 2004.

[2] N. Jacobs and R. Pless, "Time scales in video surveillance," *IEEE TCSVT*, vol. 18, no. 8, pp. 1106-1113, Aug. 2008.

[3] F.-H. Cheng and Y.-L. Chen, "Real time multiple objects tracking and identification based on discrete wavelet transform," *Pattern Recognition*, vol. 39, no. 3, pp. 1126-1139, Jun. 2006.

[4] B. Sugandi, H. Kim, J. K. Tan, and S. Ishikawa, "Tracking of moving objects by using a low resolution image," *ICICIC*, pp. 408-408, Sep. 2007.

[5] C.-H. Hsia, J.-M. Guo, and J.-S. Chiang, "An improved low complexity algorithm for 2-D integer lifting-based discrete wavelet transform using symmetric mask-based scheme," *IEEE TCSVT*, vol. 19, no 8, pp. 1201-1208, Aug. 2009.

[6] C. Kotropoulos and I. Pitas, "Constrained adaptive LMS L-filters," *IEEE ICASSP*, vol.3, no. 12, pp.1596-1609, Dec. 1996.

[7] F. E. Alsaqre and Y. Baozong, "Multiple moving objects tracking for video surveillance system," *IEEE ICSP*, vol. 2, pp.1301-1305, Aug. 2004.