Moving Object Tracking Using Symmetric Mask-Based Scheme

Chih-Hsien Hsia, Ding-Wei Huang, Jen-Shiun Chiang, and Zong-Jheng Wu
Department of Electrical Engineering, Tamkang University, Taiwan
Email: chhsia@ee.tku.edu.tw; dwhuang@ee.tku.edu.tw; s697450020@s97.tku.edu.tw

Abstract

This work presents a new approach, symmetric mask-based scheme (SMBS), for moving object detection and tracking based on the symmetric mask-based discrete wavelet transform (SMDWT). This work presents a fast algorithm, called 2-D SMDWT, to improve the critical issue of the 2-D lifting-based Discrete Wavelet Transform (LDWT), and then obtains the benefit of low latency, reduced complexity, and low transpose memory for object detection. The successful moving object detection in a real surrounding environment is a difficult task due to noise issues such as fake motion or Gaussian noise. The SMBS approach can effectively reduce noises with low computing cost in both indoor and outdoor environments. The experimental results indicate that the proposed method can provide precise moving object detection and tracking.

1. Introduction

The intelligent visual surveillance system [1][2] can detect moving objects in the initial stage and subsequently process the functions such as object classification, object tracking, and object behaviors description. Detecting moving object is a basic and significant task in every surveillance application. The accurate location of the moving object does not only provide a focus of attention for post-processing but also can reduce the redundant computation for the incorrect motion of the moving object. The successful moving object detection in a real surrounding environment is a difficult task, since there are many kinds of problems such as illumination changes, fake motion, and Gaussian noise in the background that may lead to detect incorrect motion of the moving object.

There are three typical approaches for motion detections [1]-[3]: Background subtraction, temporal differencing, and optical flow. Generally, the above three moving object detection methods are all sensitive to illumination changes, noises, and fake motion such as moving leaves of trees. In order to solve these problems, several approaches by using low resolution images for object detecting and tracking were proposed [4]-[6]. Video tracking systems have to deal with variously shaped and sized input objects, which often result in a poor computing efficiency of the input of images. Usually a low resolution image is insensitive to illumination changes and can reduce the small movement like moving leaves of trees in the background. Although these approaches can deal with small movement, the 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, symmetric mask-based scheme (SMBS), for detecting and tracking moving objects by using symmetric mask-based discrete wavelet transform (SMDWT) [7]. In SMBS, only the LL (5×5) mask band of SMDWT is used. Unlike the conventional DWT method to processing row and column dimensions separately by low-pass filter and down-sampling, the LL mask band of SMDWT is used to directly calculate the LL band image. Our proposed method can reduce the image transfer computing cost and remove fake motion that is not belonged to the real moving object. Furthermore, it can retain a better slow motion of the object than that of the low resolution method [6] and provide effective and complete moving object regions.

2. Wavelet Transform and Low Resolution Technique

Before dealing with the motion object detection, there are some methods for removing noises or fake motion and reducing computing cost proposed in the past several years. DWT [8] and low resolution technique [6] are two important approaches, and are briefly described in this section.

Discrete wavelet transform

Each subband image has its own feature, such as the low-frequency information is preserved in the LL band and the high-frequency information is almost preserved in the HH, HL, and LH bands. The LL subband image can be further decomposed in the same way for the second level subband image. By using the 2-D DWT, the image can be decomposed into any level subband images.

Cheng *et al.* [4] applied the 2-D DWT for detecting and tracking moving objects and only the LL_3 is used for detecting the moving object motion. Because noises are preserved in high-frequency, it can reduce computing cost for post-processing by using an LL_3 band image. This method can be used for coping with noise or fake motion effectively, however the traditional DWT scheme has the disadvantages of complicated calculation when an original image is decomposed into the LL band image. Moreover if it uses an LL_3 band image to deal with the fake motion it

978-0-7695-3744-3/09 $25.00 © 2009 IEEE
DOI 10.1109/IAS.2009.150

may cause incomplete moving object detecting regions.

2.2 Low resolution method

Sugandi *et al*. [6] proposed a simple method by using the low resolution concept to deal with the fake motion such as moving leaves of trees. The low resolution image is generated by replacing each pixel value of an original image with the average value of its four neighbor pixels and itself as shown in Fig. 1. It also provides a flexible multi-resolution image like the discrete wavelet transform. Nevertheless, the low resolution images generated by using the 2×2 average filter method are more blurred than that by using the DWT method. It may reduce the preciseness of post-processing (such as object tracking and object identification), because the post-processing depends on the correct location of the moving object detecting and accuracy moving object data.

3. Symmetric Mask-Based Scheme

In order to detect and track the moving object more accurately, we propose a method called symmetric mask-based scheme (SMBS) that is based on the symmetric mask-based discrete wavelet transform (SMDWT) [7]. It does not only retain the features of the flexibilities for multi-resolution, but also does not cause high computing cost when using it for finding different subband images. In addition, it preserves more image quality of the low resolution image than that of the low resolution method [6].

Symmetric mask-based discrete wavelet transform (SMDWT)

In 2-D DWT, the computation needs large transpose memory requirement and has a long critical path. On the other hand, SMDWT overcomes drawbacks of the conventional LDWT that uses integer operation and needs less transpose memory. The SMDWT has many advanced features such as short critical path, high speed operation, regular signal coding, and independent subband processing [7]. The derivation coefficient of the 2-D SMDWT is based on the 2-D 5/3 integer LDWT. For computation speed and simplicity considerations, four masks, 3×3, 5×3, 3×5, and 5×5, are used to perform spatial filtering tasks. Moreover, the four-subband processing can be further optimized to speed up and reduce the temporal memory of the DWT coefficients. The four matrix processors consist of four mask filters, and each filter is derived from one 2-D DWT of 5/3 integer lifting-based coefficients. The block diagram of the 2-D SMDWT is shown in Fig. 2.

3.1.1. Low-Low (LL) band mask coefficients reduction for 2-D SMDWT. According to the 2-D 5/3 LDWT, the LL-band coefficients of the SMDWT can be expressed as follows:

$$LL(i,j)=(9/16)x(2i,2j)+(1/64)\sum_{u=0}^{1}\sum_{v=0}^{1}x(2i-2+4u,2j-2+4v)+(1/16)\sum_{u=0}^{1}\sum_{v=0}^{1}x(2i-1+2u,2j-1+2v)+(-1/32)\sum_{u=0}^{1}\sum_{v=0}^{1}x(2i-1+2u,2j-2+4v)+(-1/32)\sum_{u=0}^{1}\sum_{v=0}^{1}x(2i-2+4u,2j-1+2v)+(3/16)\sum_{u=0}^{1}[x(2i-1+2u,2j)+P(2i,2j-1+2u)]+(-3/32)\sum_{u=0}^{1}[x(2i-2+4u,2j)+x(2i,2j-2+4u)]. \quad (1)$$

The mask as shown in Fig. 2(a) can be obtained via Eq. (1), where α=-1/32, β=1/64, γ=1/16, δ=-3/32, ε=3/16 and ζ=9/16. The DSP and hardware architecture are depicted in Fig. 2(b). The complexity of the SMDWT is further reduced by employing the symmetric feature of the mask. Thus, HH, HL, and LH-bands can be done in the same manner.

3.1.2. Performance results. This work uses the symmetric feature of the masks in SMDWT to improve the design. Experimental results, as shown in Table 1 show that the proposed algorithm is superior to most of the previous works [8]-[11]. The SMDWT approach requires a transpose memory of size (*N*/2)+26 ((*N*/2) is on-chip memory of size and 26 is the number of registers).

The flowchart of the proposed SMBS moving object detection and tracking system is shown in Fig. 3. Basically we apply the temporal differencing 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 two-level LL band frames (LL_{2t-1}, LL_{2t}, and LL_{2t+1}) by using SMDWT. After most of the noises and fake motions are moved into the high-frequency subband, it can proceed with the post-processing by employing these three two-level LL band frames. B_{t-1} and B_t can be obtained by computing the binary values of these three successive two-level LL band frames (in between LL_{2t-1}, $LL_{2t,}$ and LL_{2t+1}) and a threshold value *T* in Eq. (2).

$$B_{t-1}(i,j) = \begin{cases} 1, & \text{if } |LL2_{t-1}(i,j) - LL2_t(i,j)| > T \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

$$B_t(i,j) = \begin{cases} 1, & \text{if } |LL2_t(i,j) - LL2_{t+1}(i,j)| > T \\ 0, & \text{otherwise} \end{cases}.$$

The motion mask (MM_t) can be generated by using the union operation of B_{t-1} and B_t (using logical OR). 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 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. Then we apply erosion operator for eliminating redundant pixels in the motion mask boundary.

It scans the eight neighbors of the motion mask MMR_t 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 [4] to track the moving object motion. For this reason, we find the bounding box according to its motion mask from the foregoing work. The bounding box is made by finding the minimum and maximum values of row and column coordinators of the motion mask and connecting them. In order to track moving objects in the original image size, we have to transform the coordinator from the LL_2 image size back to the original image size according the spatial relationship of the discrete wavelet transform.

4. Experimental Results

In this work, the experimental results of several different environments including indoor and outdoor environments with statistic video system are demonstrated. The original image frame size is 320×240, and the format of a color image frame is 24-bit in a RGB system. We use all gray level frames from transferring RGB system to YCbCr system for detecting moving object motion and utilize the LL_2 image size of 80×60 generated by using SMDWT from the original image for our proposed moving object detection and tracking system.

Dealing with noise issues

Different LL band images including one-level, two-level, and three-level LL band images are used to deal with noises and compare with their results. We suggest that a successful eliminating noise image has no other motion mask besides moving object motion masks. Each level LL band image has effective results when dealing with indoor noises like Gaussian noise produced by the random noise and the statistical noise. However, when the outdoor noise such as moving leaves of trees, the LL_1 band image has poor results because these outdoor noises sometimes are large that cannot be eliminated completely.

Moving object tracking

We consider that it has a complete moving object region if it is a successful work. These moving object regions cannot have only a part of moving object, and that will be treated as a fail tracking. By experiments, without SMBS technique many noise masks are tracked. However, even if the moving objects are tracked, those moving regions are fragmented. By using SMBS, the noises can be filtered out. It still generates incomplete moving object regions by using LL_1 image, because the relevance of these pixels in the LL_1 image is deleted. When using a three-level resolution image to detect the moving objects, it generates incompletely moving object regions, owing to the LL_3 image causing too many slow motions belonged to the moving object disappeared. Finally, let us look at the results of the LL_2 band image. Using the two-level band image has a better tracking region and also can cope with noises and fake motion effectively. Table 2 shows the accuracy rate of the tracking moving objects.

We use the 2×2 average filter in substitution for the original SMBS block in our system flowchart to demonstrate the moving object that is blurred with the 2×2 average filter. Experimental results show the accuracy rate of successful object tracking with the 2×2 average filter in Table 3. It is easy to perceive the contrast between Table 2 and Table 3 of any resolution image, and the LL image generated by the SMDWT has a better successful rate than the low resolution image generated by the 2×2 average filter.

5. Conclusions

A wavelet-based scheme using SMBS for moving object detection and tracking is proposed in this paper. It is able to detect and track moving objects in indoor and outdoor environments with statistic video system. We adopt the SMDWT to replace the traditional DWT. It does not only overcome the drawbacks of high complex computation and slow speed for the traditional DWT, but also preserves the wavelet features of the flexible multi-resolution image and the capability for dealing with noises and fake motion such as moving leaves of trees. The SMDWT also has the advantages of reduced complexity, regular signal coding, short critical path, reduced latency time and independent subband coding processing. Furthermore, the 2-D LDWT performance can also be easily improved by exploiting appropriate parallel method inherently in SMDWT. The experimental results demonstrate that the 2-D 2-L LL band image can effectively track moving objects by region-based tracking under any environments, as well as it can cope with noise issues. The SMBS system can be extended to the real-time visual surveillance system applications, such as object classification and descriptive behaviors of object.

6. References

[1] W.-M. Hu, T.-N. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on System, Man, and Cybernetics*, vol. 34, no. 3, pp. 334-352, August 2004.

[2] N. Jacobs and R. Pless, "Time scales in video surveillance," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1106-1113, August 2008.

[3] 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," *Carnegie Mellon University*, Technical Report, CMU-RI-TR-00-12, 2000.

[4] 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, June 2006.

[5] F. E. Alsaqre and Y. Baozong, "Multiple moving objects tracking for video surveillance system," *IEEE International Conference on Signal Processing*, vol. 2, pp.1301-1305, August 2004.

[6] B. Sugandi , H. Kim, J. K. Tan, and S. Ishikawa, "Tracking of moving objects by using a low resolution image," *International Conference on Innovative Computing, Information and Control*, pp. 408-408, September 2007.

[7] 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 Transactions on Circuits and Systems for Video Technology*, Accepted paper, 2009.

[8] K. Andra, C. Chakrabarti, and T. Acharya, "A VLSI architecture for lifting-based forward and inverse

wavelet transform," *IEEE Transactions on Signal Processing*, vol. 50, no.4, pp. 966-977, April 2002.

[9] K.C.B. Tan and T. Arslan, "Shift-accumulator ALU centric JPEG 2000 5/3 lifting based discrete wavelet transform architecture," *IEEE International Symposium on Circuits and Systems,* vol. 5, pp. V161-V164, May 2003.

[10] H. Varshney, M. Hasan, and S. Jain, "Energy efficient novel architecture for the lifting-based discrete wavelet transform," *IET Image Process*, vol. 1, no. 3, pp.305-310, September 2007.

[11] P.-Y. Chen, "VLSI implementation of discrete wavelet transform using the 5/3 filter," *IEICE Transactions on Information and Systems*, vol. E85-D, no.12, pp. 1893-1897, December 2002.

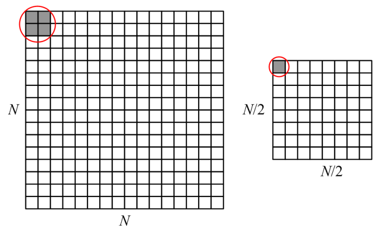


Figure 1. Diagram of the 2×2 average filter method.

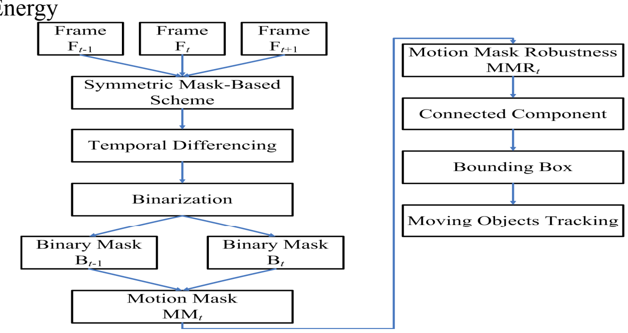
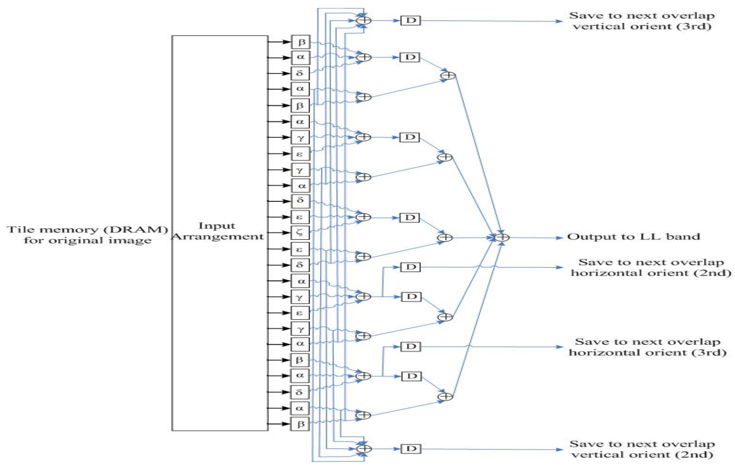


Figure 3. Flow chart of the moving object detection and tracking based on SMBS.

β	α	δ	α	β
α	γ	ε	γ	α
δ	α	ζ	α	δ
α	γ	ε	γ	α
β	α	δ	α	β

(a)



(b)

Fig. 2. LL band mask coefficients and the corresponding DSP architecture. (a) Coefficients. (b) Hardware architecture design.

Table 1. Performance comparisons.

Methods	2-D DWT	Wave stage	Transpose memory	Latency	Computing time	Complexity
[8]	LDWT	Integer	3.5N	2N+5	$(N^2/2)+N+5$	Simple
[9]	LDWT	Integer	3N	N/A	$(N^2/2)+N+5$	Medium
[10]	LDWT	Integer	3N	13	N/A	Medium
[11]	LDWT	Integer	3N	N/A	$(N^2/2)+N+5$	Medium
Proposed	SMDWT	Integer	(N/2)+26	2	$N^2/4+3$	Simple

Table 2. The accuracy rate of successful object tracking with the SMDWT.

Level	Environment	Total # of frames	Total # of moving objects	# of successful moving objects tracking	Accuracy rate
Original	Indoor	1,750	892	639	71.64%
Level = 2				828	92.83%
Original	Outdoor	2,195	1,621	1,315	81.12%
Level = 2				1,502	92.66%

Table 3. The accuracy rate of successful object tracking with the 2×2 average filter [6].

Size of image in low resolution	Environment	Total #of frames	Total # of moving objects	# of successful moving objects tracking	Accuracy rate
1/16	Indoor	1,750	892	615	68.95%
1/16	Outdoor	2,195	1,621	972	59.96%