

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Multiple Moving Objects Detection and Tracking Using Discrete Wavelet Transform

Chih-Hsien Hsia, Jen-Shiun Chiang¹ and Jing-Ming Guo

Department of Electrical Engineering

National Taiwan University of Science and Technology, Taipei

¹*Department of Electrical Engineering*

Tamkang University, Taipei

Taiwan

1. Introduction

In recent years, video surveillance systems for the purpose of security have been developed rapidly. More and more researches try to develop intelligent video surveillance systems to replace the traditional passive video surveillance systems (Hu et al., 2004) and (Jacobs & Pless, 2008). The intelligent video surveillance system 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 very important aspect of computer vision and has a very wide range of surveillance applications. 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 (Cheng & Chen, 2006), night detection (Huang, 2008), and Gaussian noise in the background (Gonzalez & Woods, 2001) that may lead to detect incorrect motion of the moving object. There are three typical approaches for motion detection (Hu et al., 2004), (Jacobs & Pless, 2008), and (Collins, 2000): background subtraction, temporal differencing, and optical flow. The background subtraction method detects moving regions between the current frame and the reference background frame. It provides the most complete motion mask data, but is susceptible to dynamic scene changes due to lighting and extraneous events. Therefore, it has to update the reference background frame frequently. The temporal differencing approach extracts the moving region by using consecutive frames of the image sequences. It is suitable for dynamic environment, but often extracts incomplete relevant motion object pixels. The optical flow method uses characteristics of flow vectors of moving objects over time to detect moving regions. However, most optical flow methods are with higher complex computation. 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 the mentioned problems, several approaches for object detecting and tracking were proposed (Ahmed et al., 2005), (Alsaqre & Baozong, 2004), (Cheng & Chen, 2006), (Chen & Yang, 2007), (Collins, 2000), (Cvetkovic et al., 2006), (Hsieh & Hsu, 2007), (Hu et al., 2004), (Hu et al., 2009), (Huang et al., 2008), (Jacobs & Pless, 2008), (Liu et al., 2006), (Mckenna, 2000), (Sugandi, 2007), and (Tab, 2007). Video tracking systems have to deal with variously shaped and sized input objects, which often result in a massive computing cost of the input of images. Cheng *et al.* (Cheng & Chen, 2006) 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). It only processes the part of LL-band image due to the consideration of low computing cost and noise reduction issues. Although this method provides low computing cost (low resolution) for post-processing and noise reduction based on the conventional DWT, 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. Especially they use the three-level low-low band image (LL_3) that does not only bring a great image size transfer computation, but also the slow motion of the real moving objects may disappear. After dealing with the background subtraction, Alsaqre *et al.* (Alsaqre & Baozong, 2004) 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.* (Sugandi et al., 2007) proposed a method for detecting and tracking objects by using a low resolution image with the 2×2 average filter (2×2 AF), which is generated by replacing each pixel value of the original image with the average value of its neighbors and itself. 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, direct LL-mask band scheme (DLLBS), for detecting and tracking moving objects by using SMDWT (Hsia et al., 2009). In DLLBS, we can select only the LL-mask band of SMDWT. Unlike the conventional DWT method to process row and column dimensions separately by low-pass filter and down-sampling, the LL-mask band of SMDWT can be 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. For objects occlusion, a new approach, characteristic point recognition (CPR), was proposed. Combined with DLLBS and CPR, it can have accurate object tracking for various types of occlusions. Furthermore, it can retain a better slow motion of objects than that of the low resolution method (Sugandi et al., 2007) and provide effective and complete moving object regions.

2. Discrete Wavelet Transform and low resolution technique

Due to the imperfection of video acquisition systems and transmission channels, images are often corrupted by noise. Therefore, this degradation leads to a significant reduction of image quality, especially for the task that performs high-level computer vision, such as object tracking, recognition, etc. Before dealing with motion object detection, there are several methods for removing noises or fake motion and reducing computing cost proposed in the past several years. DWT (Cheng & Chen, 2006) and low resolution technique (Andra et al., 2000) are two important approaches, and are briefly described in the following sub-sections.

2.1 Discrete Wavelet Transform method

Wavelet transform (Mallat, 1989) was proposed in the mid-1980s, and it has been used in various fields such as signal processing, image processing, computer vision, image compression, biochemistry medicine, etc. For image processing, it provides an extremely flexible multi-resolution image and can decompose an original image into different subband images including low- and high-frequencies. Therefore people can choose the specific resolution data or subband images upon their own demands (Hsia et al., 2009), (Mallat, 1989), (Ge et al., 2007), (Liu et al., 2006), (Ahmed et al., 2005), and (Tab et al., 2007).

A 2-D DWT of an image is illustrated in Fig. 1(a). When the original image is decomposed into four-subband images, it has to deal with row and column directions separately. First, the high-pass filter G and the low-pass filter H are exploited for each row data, and then are down-sampled by 2 to get high- and low-frequency components of the row. Next, the high- and the low-pass filters are applied again for each high- and low-frequency components of the column, and then are down-sampled by 2. By way of the above processing, the four-subband images are generated: HH, HL, LH, and LL. 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 2-D DWT, an image can be decomposed into any level subband images, as shown in Fig. 1.

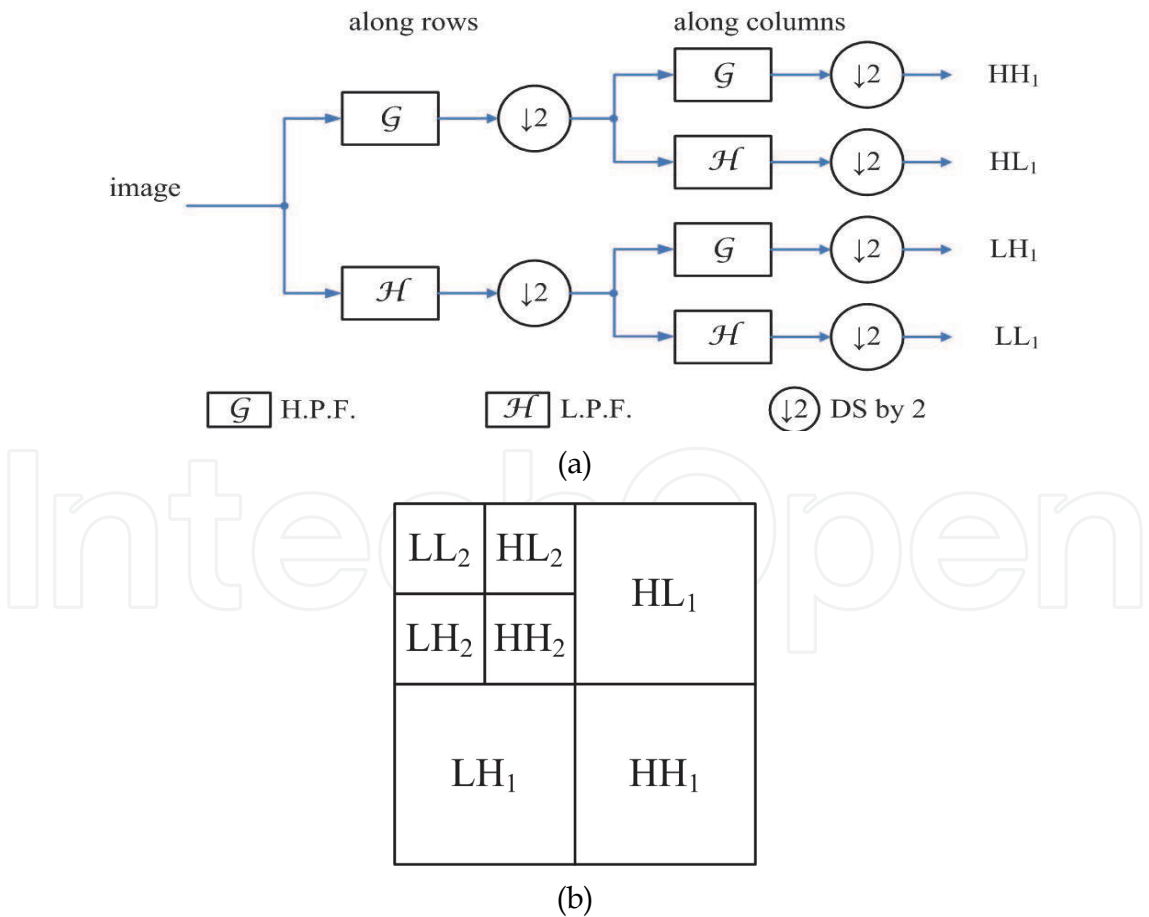


Fig. 1. Diagrams of DWT image decomposition: (a) the 1-L 2-D analysis DWT image decomposition process, (b) the 2-L 2-D analysis DWT subband.

Cheng *et al.* (Cheng & Chen, 2006) applied the 2-D DWT for detecting and tracking moving objects and only the LL_3 -band image is used for detecting the moving object motion. Because noises are preserved in high frequency, it can reduce the computing cost for post-processing by using the LL_3 -band image. This method can be used for coping with noise or fake motion effectively, however the conventional 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 may cause incomplete moving object detecting regions.

2.2 Low resolution method

Sugandi *et al.* (Sugandi, 2007) 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. 2. It also provides a flexible multi-resolution image like the DWT. Nevertheless, the low resolution images generated by using the 2×2 average filter method are more blurred than that by using the DWT method, as shown in Fig. 3. The average filtering is a low pass filter which denoises the image and performs restoration by the noise reduction spatial domain. It may reduce the preciseness of post-processing operation (such as occlusion and object identification), because the post-processing depends on the correct location of the moving object detecting and accuracy moving object data.

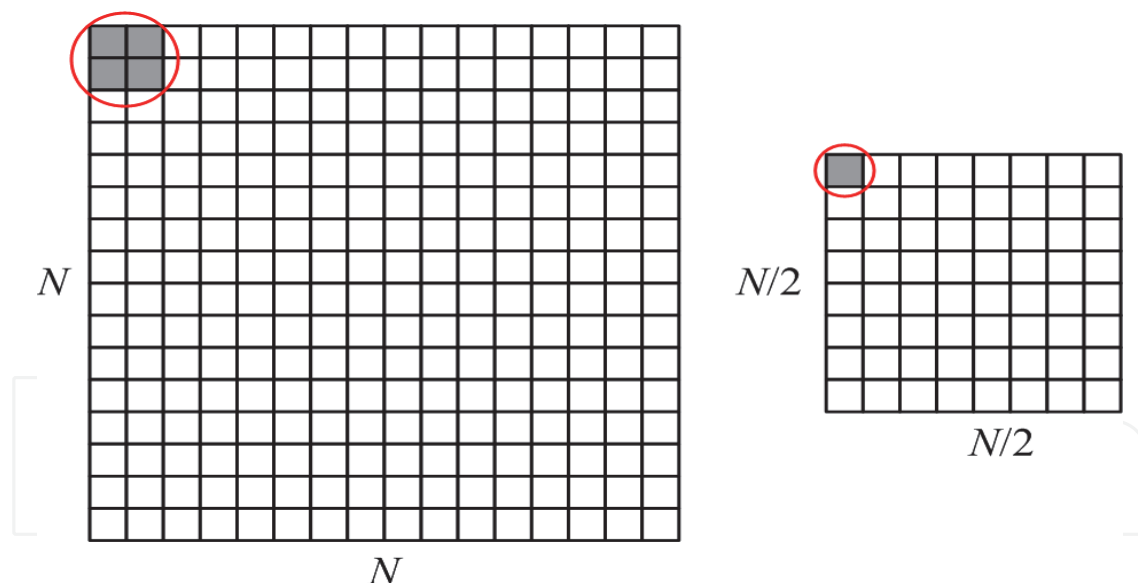


Fig. 2. Diagram of the 2×2 average filter method.

3. Direct LL-mask band scheme

In order to detect and track the moving object more accurately, we propose a new method called direct LL-mask band scheme (DLLBS) that is based on the 2-D integer symmetric mask-based discrete wavelet transform (SMDWT) (Hsia *et al.*, 2009). 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 (Sugandi, 2007).

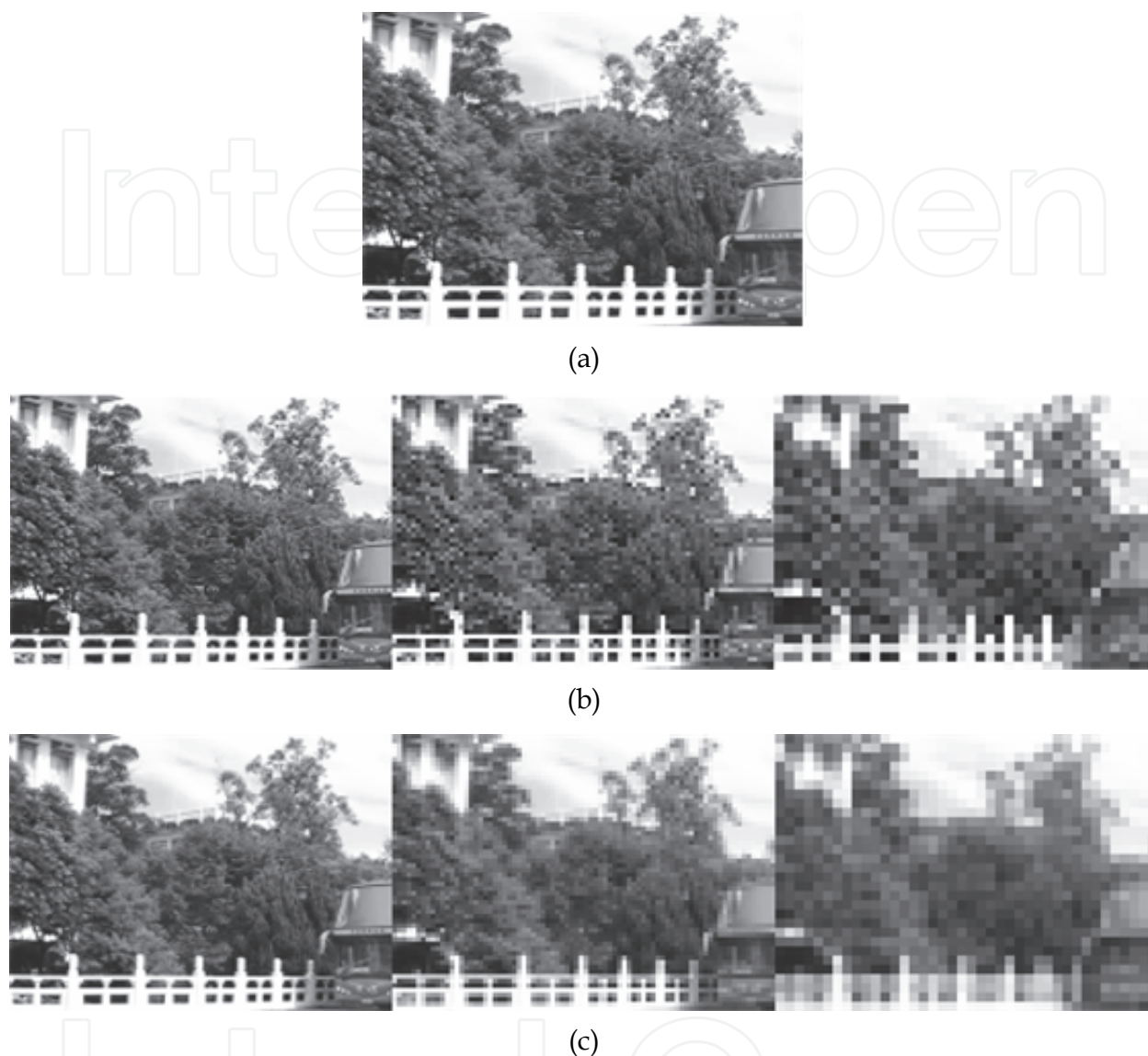


Fig. 3. Comparisons of low resolution images: (a) the original image (320×240), (b) each subband image with DWT from left to right as 160×120 , 80×60 , and 40×30 , respectively, (c) each resolution image with the 2×2 average filter method from left to right as 160×120 , 80×60 , and 40×30 , respectively.

3.1 Symmetric Mask-based Discrete Wavelet Transform (SMDWT)

In 2-D DWT, the computation needs a large transpose memory requirement and has a long critical path. The SMDWT has many advanced features such as short critical path, high speed operation, regular signal coding, and independent subband processing (Hsia et al., 2009). 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 (Hsia et al., 2009). The coefficients of each subband mask are shown in Fig. 4, and the block diagram of the 2-D SMDWT is shown in Fig. 5.

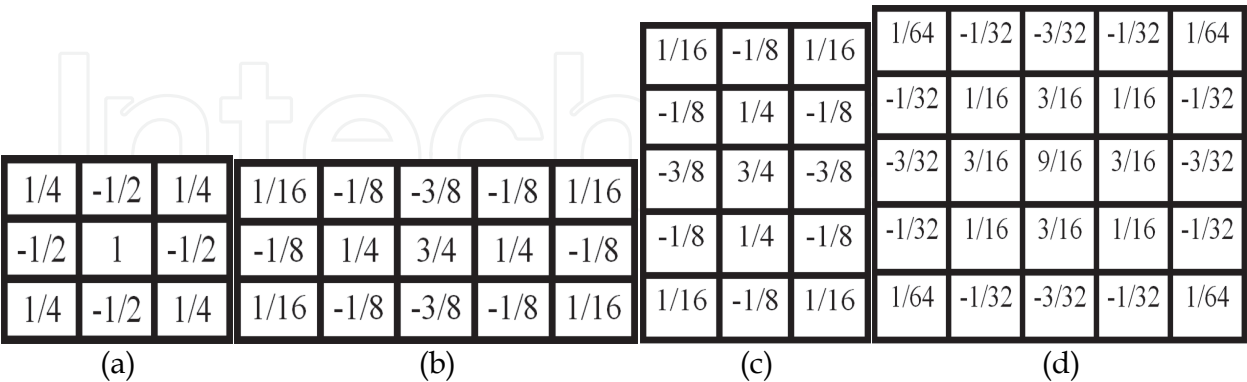


Fig. 4. The subband mask coefficients of (a) HH, (b) HL, (c) LH, and (d) LL.

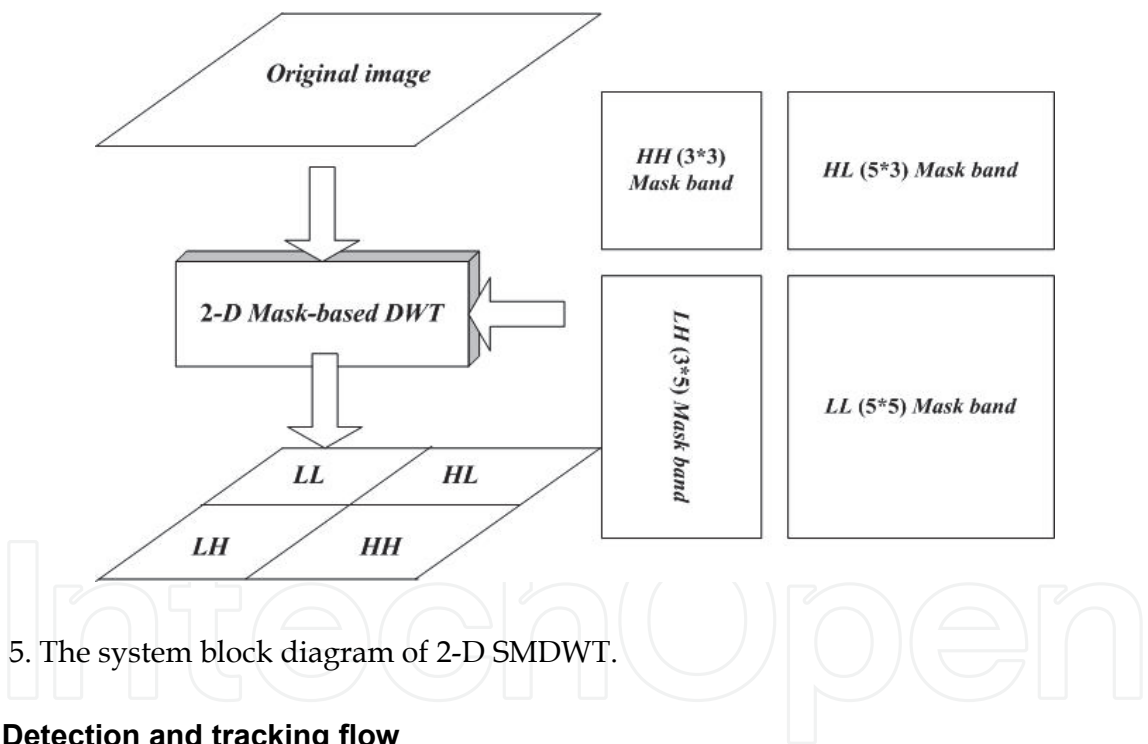


Fig. 5. The system block diagram of 2-D SMDWT.

3.2 Detection and tracking flow

The pre-processing flowchart of the proposed DLLBS moving object detection and tracking system is shown in Fig. 6. Frist, prior to color converting RGB data to YCbCr data (using Y data only). Basically we apply the double-change-detection method (Huang et al., 2004) 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 SMDWT. After most of the noises and fake motions are moved into the high-frequency subband as shown in Fig. 7, 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 LL₂-band frames (in between LL_{2t-1}, LL_{2t}, and LL_{2t+1}) and a threshold value T in (1).

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

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

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

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

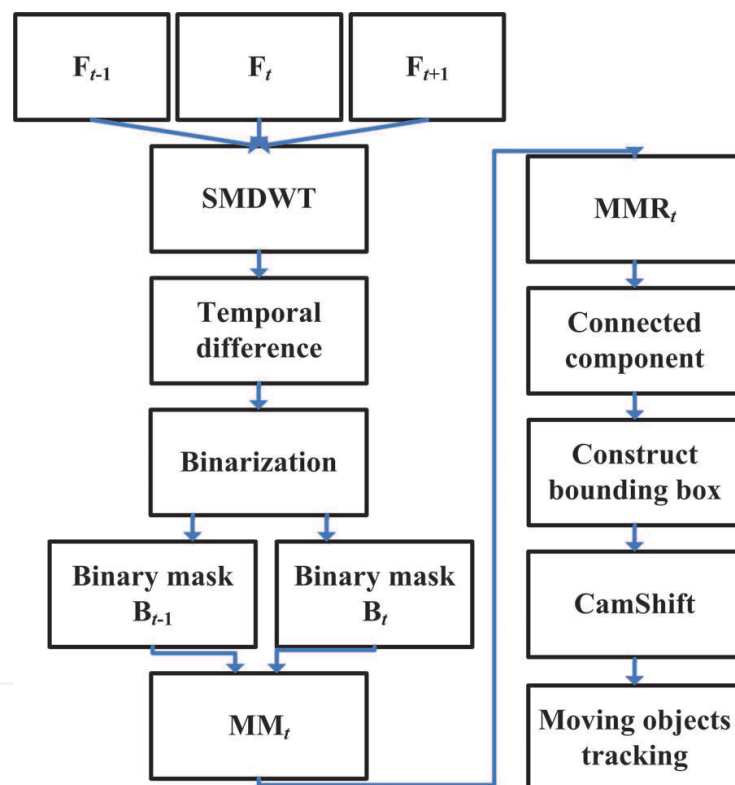


Fig. 6. The pre-processing flowchart of the moving object detection and tracking based on DLLBS.

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 (Hsieh & Hsu, 2007) 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:

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

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

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

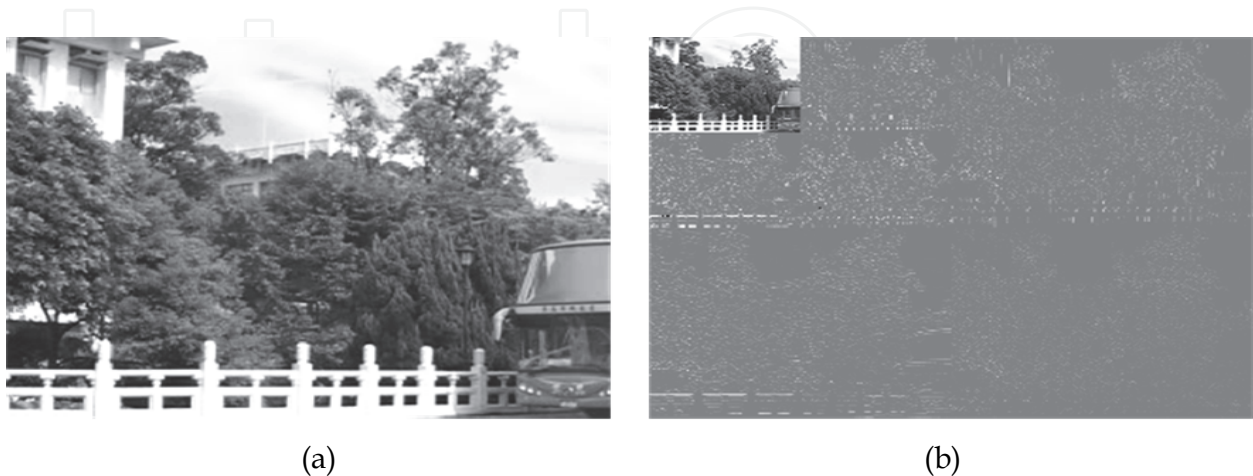


Fig. 7. After most of the noises and fake motions are removed using SMDWT (a) The original image, (b) LL_2 -band image.

It scans eight neighbors of the motion mask MMR_t image pixel by pixel from top left to bottom right (raster scan). After extracting the connected component, it obtains several moving objects. In this work, we utilize the region-based tracking algorithm (Cheng & Chen, 2006), (Mckenna, 2000), and (Chen & Yang, 2007) 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 labeling of the components based on pixel connectivity (intensity) (Gonzalez & Woods, 2001) is obtained by scanning an image and groups, pixel by pixel from top left to bottom right, in order to identify the connected pixel regions by comparing the eight neighbors that have already been encountered in the scan. If the pixel has at least one neighbor with the same label, we label this pixel as the neighbor. 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. For this reason, the bounding box is found 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 coordinates of the motion mask. 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 DWT as follows:

$$O(i,j) = LL_n(i \times 2^n, j \times 2^n). \quad (5)$$

where $n = 0 \sim l$ and l is the number of level.

In the block-matching motion estimation, the motion vector is the displacement of a block with the minimum distortion from the reference block. The CamShift block-matching algorithm determines the motion vector by identifying a block with the minimum distortion

from fast search strategies of the diamond-arc-hexagon search patterns in the search area (Chiang et al., 2008).

3.3 Occlusion handling for multiple objects tracking

In the post-processing, occlusion handling is a major problem in a video surveillance system. The most popular color space is the RGB color space (Hu et al., 2009). If the multiple objects bounding boxes are occluded, the object bounding boxes are merged into the occlusion bounding box. Here we propose a new approach for occlusion in multiple objects tracking, called characteristic point recognition (CPR). Fig. 8 shows the operation flowchart of CPR. CPR uses bounding boxes during pre-processing of DLLBS. For each tracked individual, the system will detect whether it makes occlusion with other object or group. It can obtain the RGB information from the video capture device directly to calculate the color information of the moving pixels. Owing to the information of moving pixels the size of the inter-frame difference image (1/16 of the original image) is with the central pixels.

To recognize every object, it uses the bounding box to find the characteristic point (CP). CP represents the central point of the bounding box as shown in the following equation:

$$Cs^q[n] = B_n \{(x_1, y_1), (x_2, y_2), \dots, (x_q, y_q)\} , \quad (6)$$

where $Cs^q[n]$ is an array to store the CP of every object, n the label of the object, q the amount of CP, B_n the bounding box of every object, and (x, y) the color information indexed by the position of CP. Therefore CP expresses the feature of the object. We would like to focus on each object bounding box in order to select one CP or more.

At first, the CP of every object is stored in the buffer when the first frame is input, and is regarded as the initial sample. In latter frames, the CP is matched with the sample. In other words, the CP of 1 to n matches with the CP of the sample as shown in the following equation:

$$Cd^q[n] = \text{abs}\{((Cs^q[n] - Cm^q[N])_R, (Cs^q[n] - Cm^q[N])_G, (Cs^q[n] - Cm^q[N])_B)_N\} , \quad (7)$$

where $Cm^q[N]$ is a sample array to store the CP, N the label of the sample, and $Cd^q[n]$ the absolute values obtained from the difference between $Cs^q[n]$ and $Cm^q[N]$.

After the match step, $Cd^q[n]$ stores the sample N which is identical to the object n as shown in (8):

$$L[n] = N , \quad (8)$$

where $L[n]$ is the label N to label object n . Therefore the object is recognized and labeled as sample N .

However, the objects of a frame may disappear or be occluded in latter frames. In order to hold the information of the object, the CP of the object has to be retained. Hence, we must know the object which has ever occurred when the object appears again in some frames. Because the CP may be changed by the environmental factors, the buffer has to be updated whenever a new frame is input in order to obtain the latest CP. If a new object appears, the CP of the new object should be added into the buffer to update the CP information. The CPR flowchart is shown in Fig. 8.

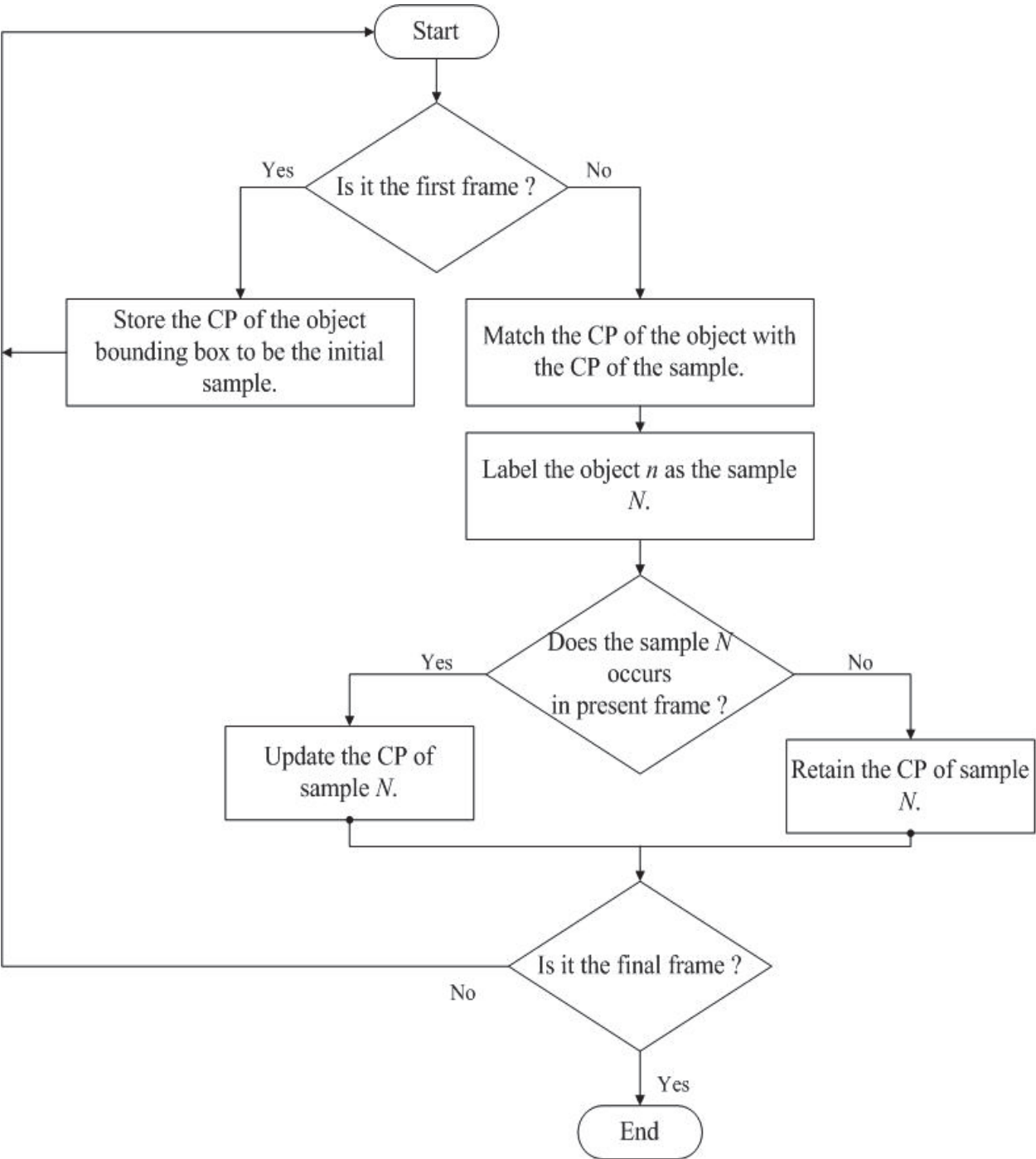


Fig. 8. CPR flowchart.

4. Experimental results

In this work, the experimental results of several different environments including indoor (all day) and outdoor (all day) environments with statistic video system are demonstrated. The original image frame sizes are 320×240, 640×480, 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 YCbCr system for detecting moving object motion and utilize the LL₂ (for 320×240) and the LL₃ (for 640×480) image size of 80×60 generated by using SMDWT from the original image for our proposed moving object detection and tracking system. The experimental environment is set using Intel 2.83 GHz Core 2 Quad CPU, 2 GB RAM, Microsoft Windows XP SP3, and Borland C++ Builder (BCB) 6.0. BCB is chosen as the software development platform. The software includes verifying for algorithms and image process for the moving objects detection.

4.1 Dealing with noise issues

There are many kinds of difficulties such as illumination changes, fake motion, and Gaussian noise in the background. Different LL-band images including one-level, two-level, three-level, and multi-level LL-band images are used to deal with noises and compare their results. We suggest that a successful eliminating noise image has no other motion mask besides moving object motion masks, as shown in Figs. 9 and 10. Table 1 shows the average (Figs. 9 and 10) successful eliminating noise rate of each level LL-band image. The first row is in the indoor environment and the second row in the outdoor environment. Each level LL-band image has effective results when dealing with indoor noises like Gaussian noise produced by random noise and statistical noise. However, when dealing with the outdoor noise such as moving leaves of trees, the LL₁-band image has poor results because these outdoor noises sometimes are large that cannot be eliminated completely.

Resolution	¹ DLLBS	² LS	³ 2×2 AFS	⁴ DSS
Level	Accuracy rate	Accuracy rate	Accuracy rate	Accuracy rate
LL ₁ (160×120)	99.54 %	99.54 %	99.07 %	98.15 %
LL ₂ (80×60)	99.07 %	99.07 %	93.07 %	81.94 %
LL ₃ (40×30)	95.83 %	95.83 %	86.11 %	63.89 %

¹DLLBS: Direct LL-mask Band Scheme; ²LS: Lifting Scheme; ³2×2 AFS: 2×2 Average Filter Scheme; ⁴DSS: Down-Sampled Scheme; ⁵Accuracy rate: Success Tracking/ Original Sequency.

Table 1. The moving objects detection and tracking results.

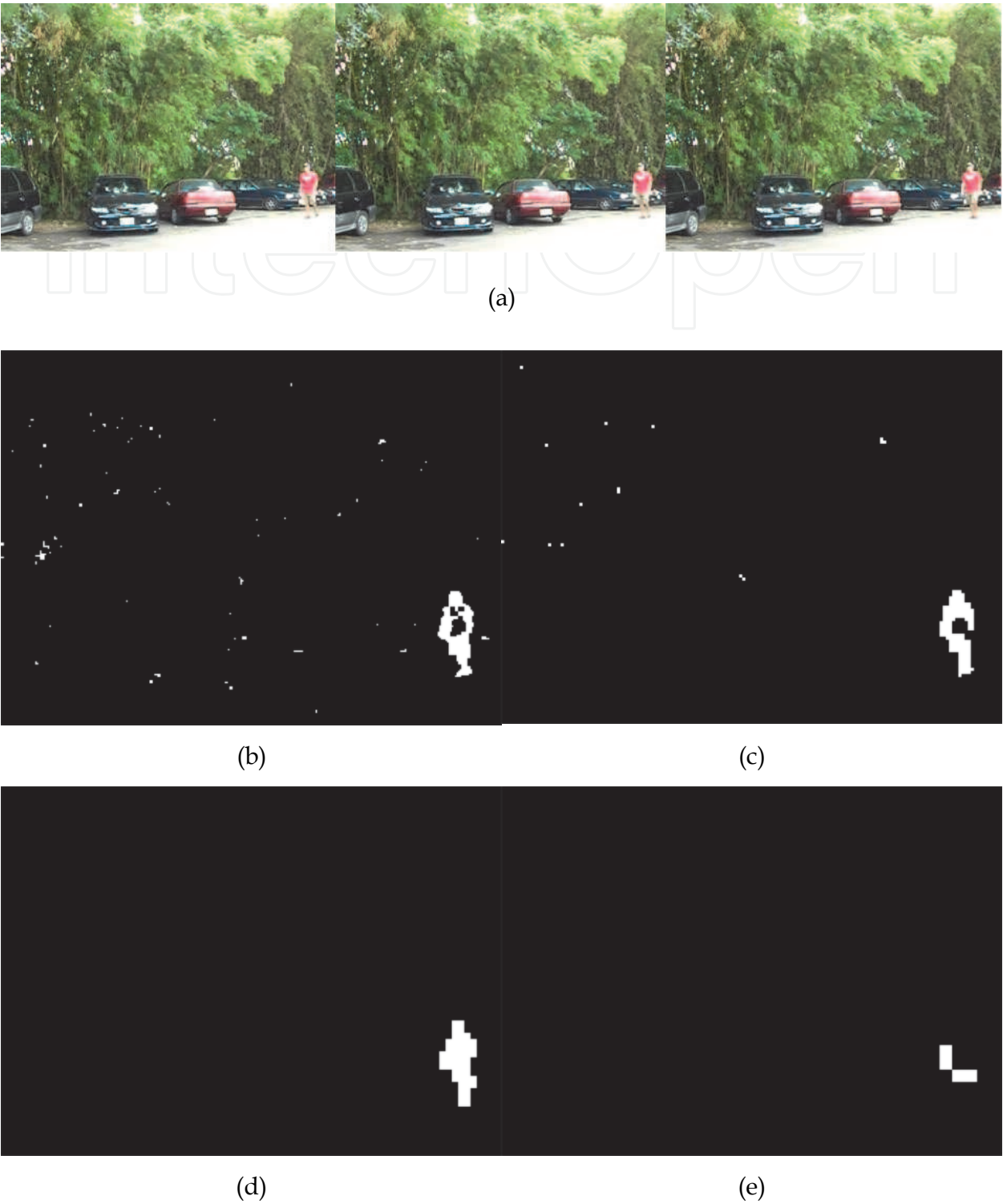


Fig. 9. Moving object detection in the outdoor environment with fake motion: (a) the original image of three consecutive frames, (b) the temporal differencing results of the original image, (c) the temporal differencing results of the LL_1 -band image, (d) the temporal differencing results of the LL_2 -band image, (e) the temporal differencing results of the LL_3 -band image.

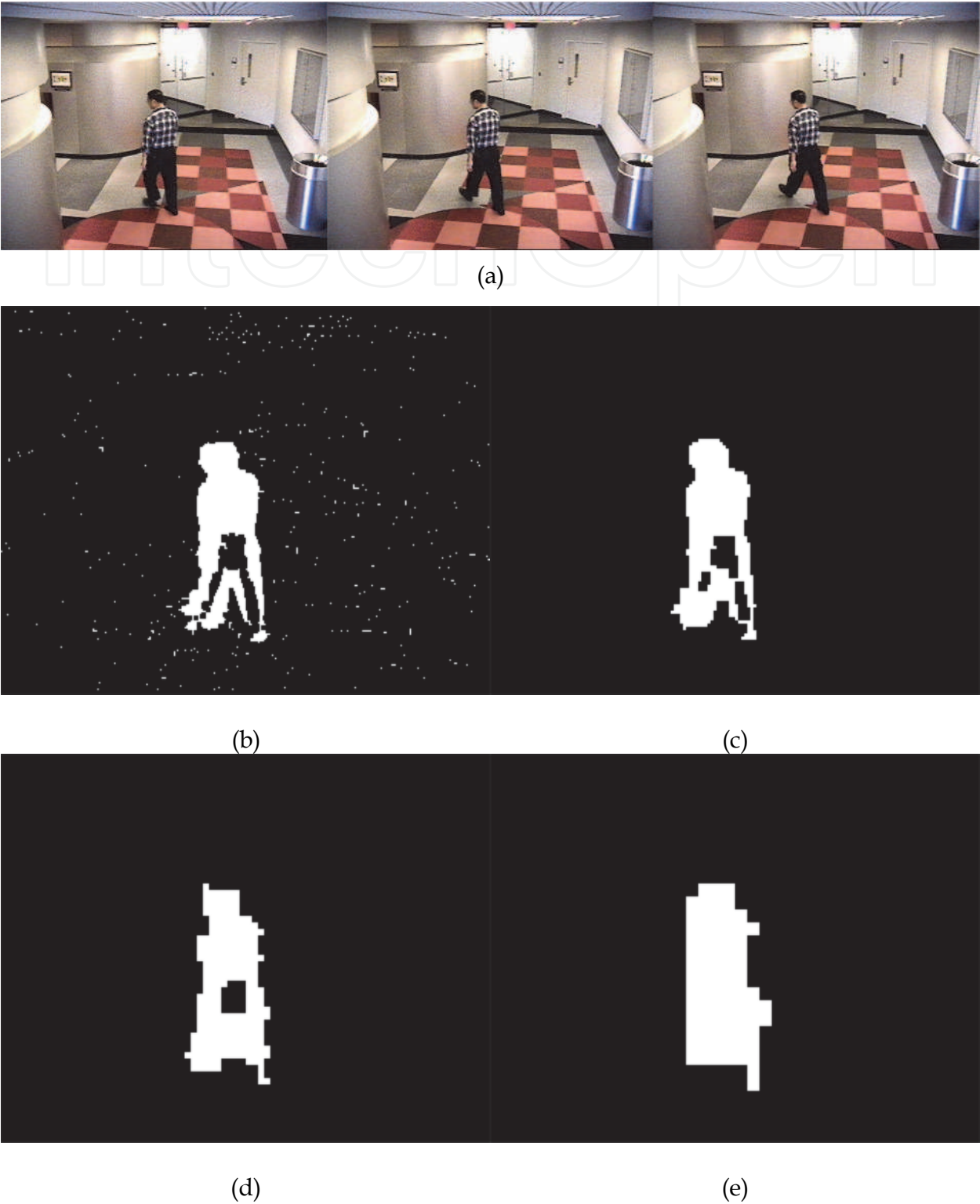


Fig. 10. Moving object detection in the indoor environment with Gaussian noise (Gonzalez & Woods, 2001): (a) the original image of three consecutive frames, (b) the temporal differencing results of the original image, (c) the temporal differencing results of the LL₁-band image, (d) the temporal differencing results of the LL₂-band image, (e) the temporal differencing results of the LL₃-band image.

4.2 Moving object tracking

We consider it to have a complete moving object region if it is a successful work, as shown in Fig. 11(a). In Fig. 11(b), the moving object regions have only a part of moving object, and that will be treated as a failure tracking. Fig. 12(a) expresses the original frame without detecting and tracking moving objects. Without DLLBS technique many noise masks are tracked. However, even if the moving objects are tracked, those moving regions are fragmented, as shown in Fig. 12(b). By using DLLBS, the noises can be filtered out, as shown in Fig. 12(c). It still generates incomplete moving object regions by using LL_1 -band image, because the relevance of these pixels in the LL_1 -band 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 -band image causing too many slow motions belonged to the moving object disappeared, as shown in Fig. 12(e). Finally, let us look at the results of the LL_2 -band image in Fig. 12(d). Using the two-level band image has a better tracking region and also can cope with noises and fake motion effectively, as shown in Table 1.



Fig. 11. Examples of (a) successful moving object tracking and (b) failure moving object tracking.

We use the 2×2 average filter scheme (AFS) in substituting the original DLLBS block system to demonstrate the moving object, however it is more blurred than the DLLBS technique. The accuracy rate of the successful object tracking with the 2×2 AFS are shown in Tables 1, 4, and 5. It is easy to perceive the contrasts between Tables 1 and 4 of any resolution image; the LL-band image generated by the DLLBS has a better successful ratio than the low resolution image generated by the 2×2 AFS.

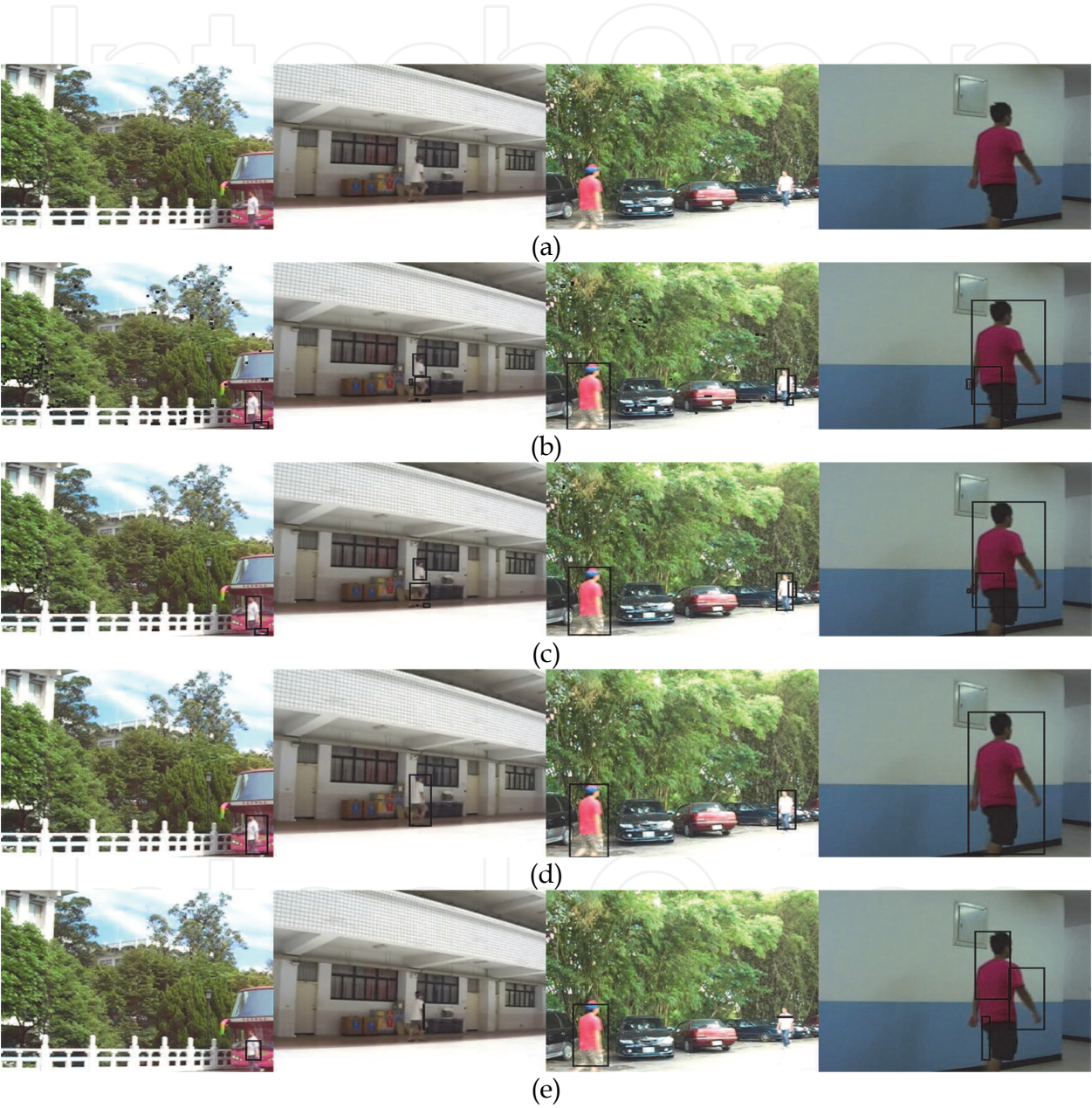


Fig. 12. Results of tracking moving objects in various environments: (a) original frames without region-based object tracking, (b) original frames with region-based object tracking, (c) LL₁-band frames with region-based object tracking, (d) LL₂-band frames with region-based object tracking, (e) LL₃-band frames with region-based object tracking.

Several experiments have been made to prove the feasibility of the proposed approach for moving object detection, tracking, and occlusion. We used an entry-level video camera and

capture card to capture the test sequences in our campus (Tamkang University), and simulated several cases of condition for moving objects, such as signal object in day time (indoor/outdoor), signal object at night (outdoor), and multiple objects in day time (outdoor) environments. All the test sequences are stored as the Microsoft AVI format with raw file of resolution 320×240, 640×480, and frame rate of 30 fps as shown in Fig. 13.

The choice of the threshold T is important. A too large value may lead to real targets missing; on the other hand a too small value may lead to noise binary images with pseudo features. The best threshold value also varies with the instantaneous illumination level. The threshold values of the best performance in different environments and DLLBS are listed in Table 2. According to the experiments, under day and midday time in outdoor for LL_2 , the threshold values of 10, 14, and 16 were applied, some noisy pixels appeared. However, when the threshold value of 15 was applied, it outperformed all other threshold values.

Resolution	Night in the outdoor	Day and Night in the indoor	Day and Midday in the outdoor
$LL_1(160\times120)$	$T = 14$	$T = 20$	$T = 25$
$LL_2(80\times60)$	$T = 4$	$T = 10$	$T = 15$
$LL_3(40\times30)$	$T = 1$	$T = 6$	$T = 11$

Table 2. The best threshold values, T , in different environments and DLLBS.

Environments	Fake motions	Low contrast	Reflection
DLLBS	Excellent	Excellent	Excellent
2×2 AFS	Excellent	Poor	Good
DSS	Poor	Excellent	Poor

Table 3. Features of various methods.

We established 16 test sequences at Tamkang University in different environments, such as day time, night time, rainy day, fast movement, slow movement, and occlusion, as shown in Fig. 13. Compared with other approaches (2×2 AFS and DSS), the DLLBS can obtain a good sparsity for spatially localized details, such as edges and singularities, as shown in Table 3. Because such details are typically abundant in natural images and convey a significant part of the information embedded therein, DWT has found a significant application for image denoising. From Tables 4, 5, and 6, we notice that some objects are not correctly identified in the test frame of the sequences. The wrong identification occurs in two reasons:

1. The moving object just enters or leaves the scene.

Because the moving object is detected and tracked at the border of the scene, the extracted features of the moving object in the case cannot represent the moving object very well.

2. The moving object is slowing down.
In this issue, the temporal difference image of the object becomes smaller and loses its situation.

Pattern	DLLBS		2×2 AFS		DSS	
	Accuracy rate	Detection + Tracking	Accuracy rate	Detection + Tracking	Accuracy rate	Detection + Tracking
Sequence1	98.61 %	53.8 FPS	98.61 %	58.5 FPS	71.76 %	52.1 FPS
Sequence2	95.90 %	56.5 FPS	96.31 %	57.1 FPS	56.15 %	49.0 FPS
Sequence3	97.40 %	54.1 FPS	92.36 %	60.7 FPS	96.59 %	63.1 FPS
Sequence4	93.55 %	54.3 FPS	82.61 %	62.0 FPS	92.65 %	63.2 FPS
Sequence5	82.97 %	56.7 FPS	80.44 %	60.2 FPS	77.29 %	65.6 FPS
Sequence6	82.57 %	53.9 FPS	78.90 %	61.5 FPS	46.79 %	63.9 FPS
Sequence7	90.28 %	54.7 FPS	37.50 %	61.4 FPS	40.28 %	63.5 FPS
Sequence8	83.33 %	55.1 FPS	73.46 %	62.4 FPS	78.40 %	61.2 FPS
Sequence9	90.16 %	53.5 FPS	75.13 %	59.0 FPS	83.94 %	58.5 FPS
Average	90.53 %	54.7 FPS	79.48 %	60.3 FPS	71.53 %	60.1 FPS

Table 4. Single moving object processing (without occlusion).

Pattern	DLLBS		2×2 AFS		DSS	
	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion
Sequence10	92.94 %	56.7 FPS	88.61 %	60.6 FPS	82.92 %	57.6 FPS
Sequence11	89.98 %	54.8 FPS	83.67 %	60.9 FPS	92.60 %	58.7 FPS
Sequence12	90.43 %	53.9 FPS	79.79 %	60.8 FPS	82.98 %	55.9 FPS
Sequence13	90.00 %	52.6 FPS	86.67 %	59.7 FPS	75.56 %	63.7 FPS
Average	90.84 %	54.5 FPS	84.69 %	60.5 FPS	83.52 %	58.9 FPS

Table 5. Multiple moving objects processing (with occlusion).

Pattern	DLLBS		2×2 AFS		DSS	
	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion
Sequence14	85.60 %	14.1 FPS	73.60 %	16.5 FPS	31.20 %	15.5 FPS
Sequence15	88.36 %	14.1 FPS	79.45 %	16.6 FPS	63.01 %	15.3 FPS
Sequence16	81.33 %	14.2 FPS	70.67 %	16.6 FPS	33.33 %	15.2 FPS
Average	85.10 %	14.1 FPS	74.57 %	16.6 FPS	42.51 %	15.3 FPS

Table 6. Multiple moving objects processing (with occlusion).



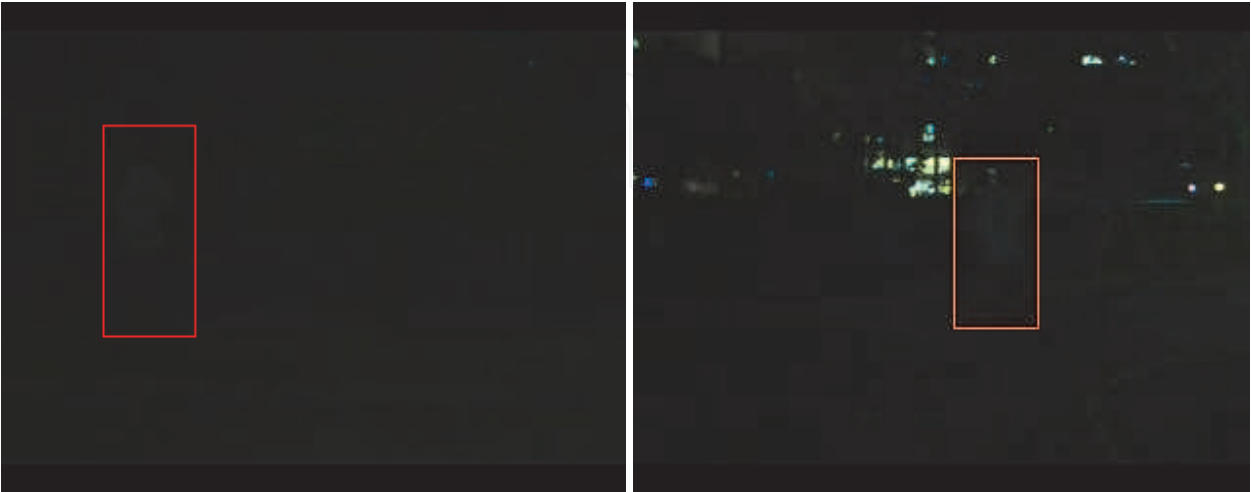
(a)

(b)



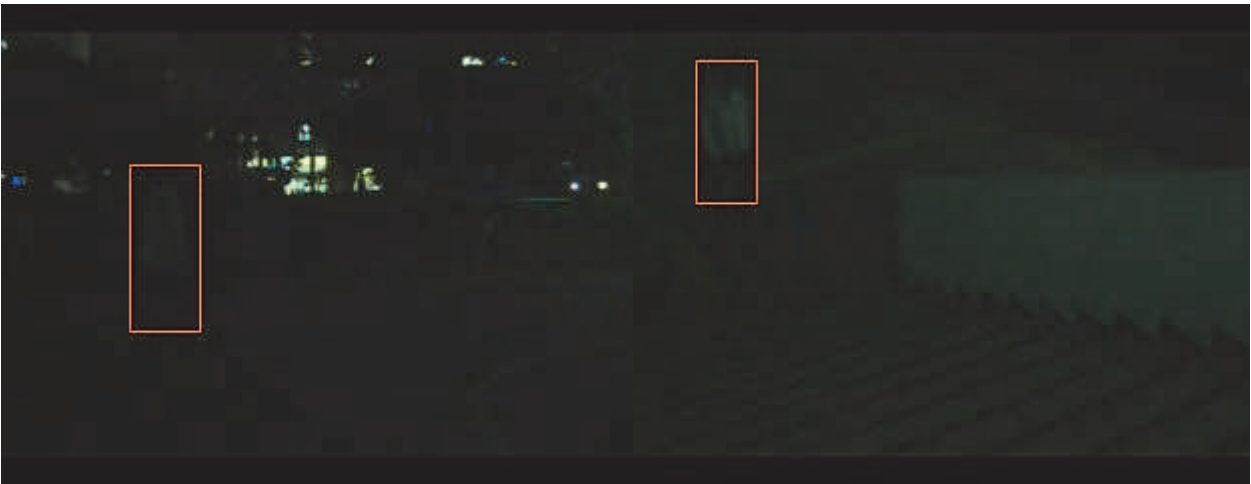
(c)

(d)



(e)

(f)



(g)

(h)



(i)

(j)



(k)

(l)

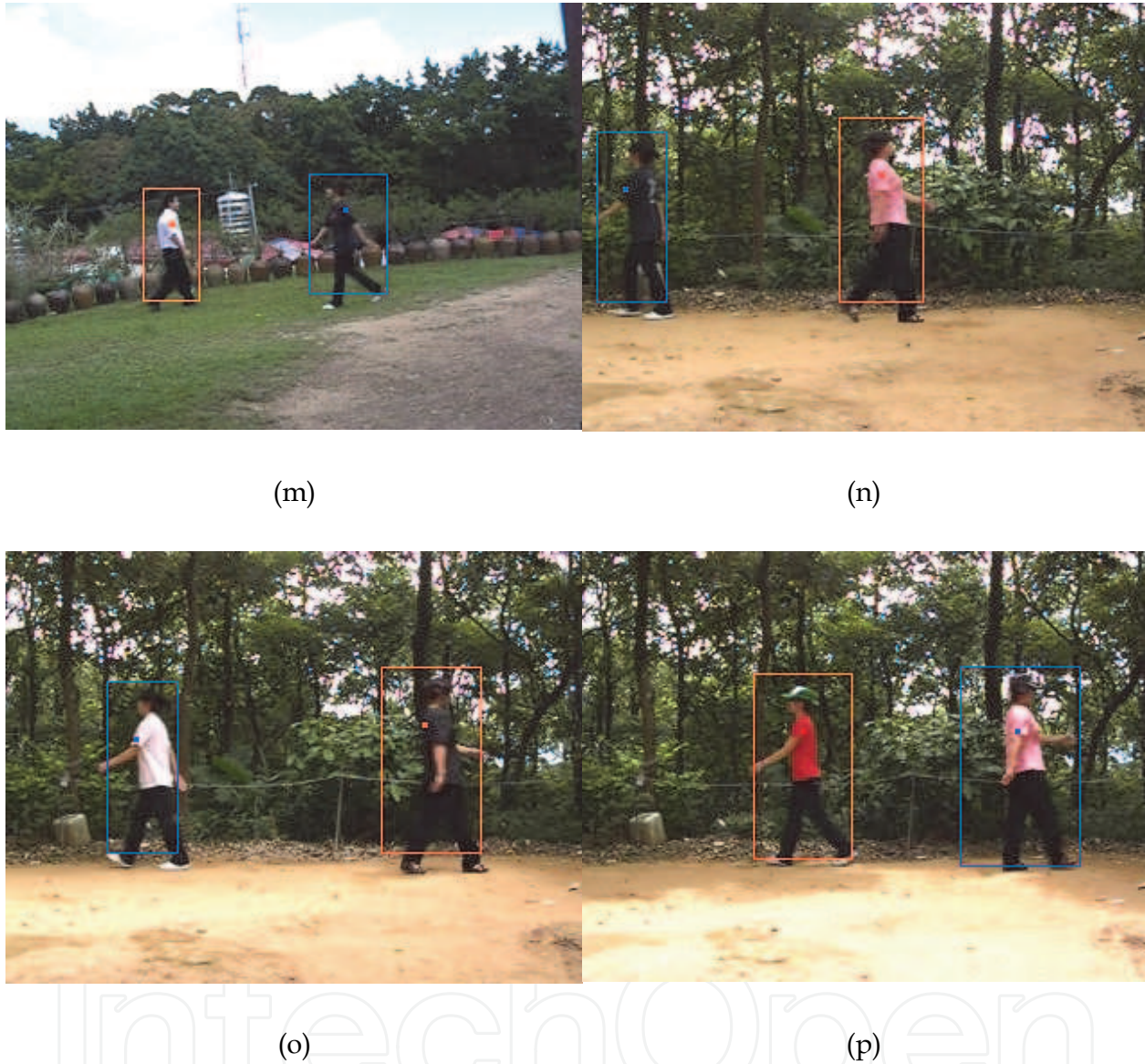


Fig. 13. Test sequences at Tamkang University: (a)-(k) are sequences 1-13 (320×240) and 14-16 (640×480); (a)-(c) show the single moving object in the outdoor; (d) single moving object in the indoor; (e) single moving object in the outdoor (fast movement to slow movement); (f) single moving object in the outdoor (fast movement); (g) single moving object in the outdoor (slow movement); (h) single moving object in the indoor (zoom-out to zoom-in); (i) single moving object in the outdoor (rainy day); (j) multiple moving object in the outdoor (occlusion); (k) multiple moving object in the outdoor; (l) multiple moving object in the outdoor (occlusion); (m) multiple moving object in the outdoor (occlusion); (n) multiple moving object in the outdoor (occlusion); (o) multiple moving object in the outdoor (occlusion); (p) multiple moving object in the outdoor (occlusion).

5. Conclusions

The direct LL-mask band scheme (DLLBS) for moving object detection and tracking is proposed in this work. It is able to detect and track moving objects in indoor and outdoor environments with statistic video systems. The proposed DLLBS 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 motion such as moving leaves of trees. In the real-word application, the experimental results demonstrate that the 2-D LL_2 -band (for 320×240) and the 2-D LL_3 -band (for 640×480) can effectively track moving objects by region-based tracking under any environments (day and night), as well as it can cope with noise issues. For occlusion considerations, we propose a new approach, characteristic point recognition (CPR). Combined with DLLBS and CPR, it can accurately track various types of occlusions. The DLLBS system can be extended to the real-time video surveillance system applications, such as object classification and descriptive behaviors of objects.

6. Reference

- Ahmed, J.; Jafri, M. N. & Ahmad, J. (2005). Target tracking in an image sequence using wavelet features and neural network, *IEEE TENCON*, (November 2005) pp.1-6.
- Andra, K.; Chakrabarti, C. & Acharya, T. (2000). A VLSI architecture for lifting-based wavelet transform, *IEEE Workshop on Signal Processing Systems*, (October 2000) pp. 70-79.
- Alsaqre, F. E. & Baozong, Y. (2004). Multiple moving objects tracking for video surveillance system, *IEEE International Conference on Signal Processing*, Vol. 2, (August 2004) pp.1301-1305.
- Chen, D.-T. & Yang, J. (2007). Robust object tracking via online dynamic spatial bias appearance models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 12, (December 2007) pp. 2157-2169.
- Cheng, F.-H. & Chen, Y.-L. (2006). Real time multiple objects tracking and identification based on discrete wavelet transform, *Pattern Recognition*, Vol. 39, No. 3, (June 2006) pp. 1126-1139.
- Chiang, J.-S.; Lin, H.-T. & Hsia, C.-H. (2008). Novel fast block motion estimation using diamond-arc-hexagon search patterns, *Journal of the Chinese Institute of Engineers*, Vol. 31, No. 6, (September 2008) pp. 955-966.
- Collins, R. T.; Lipton, A. J.; Kanade, T.; Fujiyoshi, H.; Duggins, D.; Tsin, Y.; Tolliver, D.; Enomoto, N.; Hasegawa, O.; Burt, P. & Wixson, L. (2000). A system for video surveillance and monitoring, *Carnegie Mellon University, Technical Report*, (2000) CMU-RI-TR-00-12.
- Cvetkovic, S.; Bakker, P.; Schirris, J. & With, P. H. N. de. (2006). Background estimation and adaptation model with light-change removal for heavily down-sampled video surveillance signals, *IEEE International Conference on Image Processing*, (October 2006) pp. 1829-1832.
- Daubechies, I. & Sweldens, W. (1998). Factoring wavelet transforms into lifting steps, *The Journal of Fourier Analysis and Applications*, Vol. 4, No.3, (1998) pp. 247-269.

- Ge, W.; Gao, L.-Q. & Sun, Q. (2007). A method of multi-scale edge detection based on lifting scheme and fusion rule, *International Conference on Wavelet Analysis and Pattern Recognition*, Vol. 2, (November 2007) pp. 952-955.
- Gonzalez, R. C. & Woods, R. E. (2001). *Digital image processing*, Addison-Wesley Longman Publish Co., Inc., Boston.
- Hsia, C.-H.; Guo, J.-M. & Chiang, J.-S. (2009). 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*, Vol. 19, No 8, (August 2009) pp. 1201-1208.
- Hsia, C.-H.; Guo, J.-M.; Chiang, J.-S. & Lin, C.-H. (2009). A novel fast algorithm based on SMDWT for image applications, *IEEE International Symposium on Circuits and Systems*, (May 2009) pp. 762-765.
- Hsieh, C.-C. & Hsu, S.-S. (2007). A simple and fast surveillance system for human tracking and behavior analysis, *IEEE Conference on Signal-Image Technologies and Internet-Based System*, (December 2007) pp. 812-828.
- Hu, W.-M.; Zhou, X.; Hu, M. & Maybank, S. (2009). Occlusion reasoning for tracking multiple people, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 19, No. 1, (January 2009) pp. 114-121.
- Hu, W.-M.; Tan, T.-N.; W, L. & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors, *IEEE Transactions on Systems, Man, and Cybernetics- Part C: Applications and Reviews*, Vol. 34, No. 3, (August 2004) pp. 334-352.
- Huang, K.-Q.; Wang, L.-S.; Tan, T.-I. & Maybank, S. (2008). A real-time objects detecting and tracking system for outdoor night surveillance, *Pattern Recognition*, Vol. 41, No. 1, (January 2008) pp. 423-444.
- Huang, J.-C.; Su, T.-S.; Wang, L.-J. & Hsieh, W.-S. (2004). Double-change-detection method for wavelet-based moving-object segmentation, *IET Electronics Letters*, Vol. 40, No. 13, (June 2004) pp. 798-799.
- Jacobs, N. & Pless, R. (2008). Time scales in video surveillance, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 8, (August 2008) pp. 1106-1113.
- Kharate, G. K.; Patil, V. H. & Bhale, N. L. (2007). Selection of mother wavelet for image compression on basis of nature of image, *Journal of Multimedia*, Vol. 2, No. 6, (November 2007) pp. 44-51.
- Liu, H.-H.; Chen, X.-H.; Chen, Y.-G. & Xie, C.-S. (2006). Double change detection method for moving-object segmentation based on clustering, *IEEE International Symposium on Circuits and Systems*, (May 2006) pp. 5027-5030.
- Mallat, S. G. (1989). A theory for multi-resolution signal decomposition: The wavelet representation, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, (July 1989) pp. 674-693.
- Mckenna, S. J. (2000). Tracking groups of people, *Computer Vision and Image Understanding*, Vol. 80, No. 1, (October 2000) pp. 42-56.
- Sugandi, B.; Kim, H.; Tan, J. K. & Ishikawa, S. (2007). Tracking of moving objects by using a low resolution image, *International Conference on Innovative Computing, Information and Control*, (September 2007) pp. 408-408.

Tab, F. A.; Naghdy, G. & Mertins, A. (2007). Multiresolution video object extraction fitted to scalable wavelet-based object coding, *IET Image Processing*, Vol. 1, No. 1, (March 2007) pp. 21-38.

IntechOpen

IntechOpen



Discrete Wavelet Transforms - Biomedical Applications

Edited by Prof. Hannu Olkkonen

ISBN 978-953-307-654-6

Hard cover, 366 pages

Publisher InTech

Published online 12, September, 2011

Published in print edition September, 2011

The discrete wavelet transform (DWT) algorithms have a firm position in processing of signals in several areas of research and industry. As DWT provides both octave-scale frequency and spatial timing of the analyzed signal, it is constantly used to solve and treat more and more advanced problems. The present book: Discrete Wavelet Transforms - Biomedical Applications reviews the recent progress in discrete wavelet transform algorithms and applications. The book reviews the recent progress in DWT algorithms for biomedical applications. The book covers a wide range of architectures (e.g. lifting, shift invariance, multi-scale analysis) for constructing DWTs. The book chapters are organized into four major parts. Part I describes the progress in implementations of the DWT algorithms in biomedical signal analysis. Applications include compression and filtering of biomedical signals, DWT based selection of salient EEG frequency band, shift invariant DWTs for multiscale analysis and DWT assisted heart sound analysis. Part II addresses speech analysis, modeling and understanding of speech and speaker recognition. Part III focuses biosensor applications such as calibration of enzymatic sensors, multiscale analysis of wireless capsule endoscopy recordings, DWT assisted electronic nose analysis and optical fibre sensor analyses. Finally, Part IV describes DWT algorithms for tools in identification and diagnostics: identification based on hand geometry, identification of species groupings, object detection and tracking, DWT signatures and diagnostics for assessment of ICU agitation-sedation controllers and DWT based diagnostics of power transformers. The chapters of the present book consist of both tutorial and highly advanced material. Therefore, the book is intended to be a reference text for graduate students and researchers to obtain state-of-the-art knowledge on specific applications.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Chih-Hsien Hsia, Jen-Shiun Chiang and Jing-Ming Guo (2011). Multiple Moving Objects Detection and Tracking Using Discrete Wavelet Transform, Discrete Wavelet Transforms - Biomedical Applications, Prof. Hannu Olkkonen (Ed.), ISBN: 978-953-307-654-6, InTech, Available from:
<http://www.intechopen.com/books/discrete-wavelet-transforms-biomedical-applications/multiple-moving-objects-detection-and-tracking-using-discrete-wavelet-transform>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai

www.intechopen.com

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen

IntechOpen

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen