ABSTRACT
This paper proposes wavelet-based moving object segmentation using change detection and background registration technique. The inter-frame difference for detecting change was calculated by using high frequency coefficients in wavelet subbands. The background was constructed with low frequency wavelet coefficients which were the approximated version of the original image. Combining detected changes and constructed backgrounds, we propose an efficient moving object segmentation algorithm. Our experimental results show that our proposed method obtains accurate video object planes (VOPs) in a video sequence.

KEY WORDS
Moving object segmentation, wavelet coefficient, change detection, background registration, video object plane

1. Introduction
Moving object segmentation from a video sequence is a key operation for content-based video coding [1] and multimedia description. The MPEG-4 standard provides a content-based framework to achieve the goals of fast transmission, efficient storage, flexible manipulation and the reuse of visual content. However, the MPEG-4 does not provide specific techniques for VOP (video object plane) extraction. Nonetheless, it is an indispensable process for many digital video applications. Thus, the preprocessing used to decompose sequences into VOPs is an important issue.

For VOPs extraction, object segmentation is one of the popular approaches. One of all video segmentation methods is to use spatial homogeneity [2-3] as the primary segmentation criterion. This algorithm is that the sequence is partitioned into some regions by mathematical morphology filters, and then the watershed algorithm is applied for region boundary decision. The result of this method can obtain the object boundary more precisely than other methods. However, the computation complexity is very high because both computationally intensive operation.

The change detection method for inter-frame difference [4-7, 10] is a popular scheme because it is straightforward in terms of implementation and enables the automatic detection of new appearances. The position and shape of the moving object is detected from the frame difference of two consecutive frames. The conventional change detection-based approaches [4-5, 10] are reliable for video sequences like the head-and-shoulder type of video. However, it has several drawbacks. First, the value of the frame difference depends on the speed of the object, so the quality of the segmentation cannot be maintained consistently if the speed of the object changes significantly in the sequence. It has a relatively poor result for video sequences like the surveillance type of video. Second, the uncovered background region [8] is detected as a moving object region from the frame difference information. To overcome these problems, an algorithm using the background registration technique [10] is reported. However, it is difficult to construct the background if the object motion is minute.

Change detection for moving object segmentation is also achieved in the wavelet domain [6-7]. This method may reduce noise disturbance and speed change, but the segmenting coherence is still not ensured [9].

In this paper, we propose an efficiently wavelet-based moving object segmentation algorithm that simultaneously uses the inter-frame difference and background registration technique.

The rest of this paper is organized as follows. In the next section, the proposed method is described in detail, followed by experimental results are shown in section 3. Finally, section 4 concludes this paper.

2. Proposed method
The wavelet transform decomposes an image or a video frame into high frequency and low frequency components. Since the high frequency coefficients represent the edges of an image, changes can be detected by using these coefficients. On the other hand, the low frequency components, which are an approximated version of the original signal, can be used as information for background registration. We have proposed a method for moving object segmentation using the change detection method [4-7] and background registration technique [10] in the wavelet domain.
2.1 Inter-frame difference in wavelet domain

Fig. 1 shows the overall block diagram of the proposed method. Firstly, we performed a wavelet transformation to obtain four wavelet subbands. For the high frequency subband (LH, HL, and HH), we applied the change detection method to find the change detection mask, \( CDM_{n,d}(i,j) \) as follows:

\[
CDM_{n,d}(i,j) = \begin{cases} 
1 & \text{if } |w_{n,d}(i,j) - w_{n-1,d}(i,j)| > V_{\text{th},d} \\
0 & \text{otherwise} 
\end{cases}
\]

where \( w_{n,d}(i,j) \) is the wavelet coefficient at location \((i,j)\) of frame \(n\) in each wavelet subband, and \( V_{\text{th},d} \) is a threshold determined by the significant test [11]. If the difference between inter-frame coefficients is larger than the given threshold, then we can guess that there are moving objects. However, because the extracted \( CDM_{n,d}(i,j) \) is the difference value for the edges, the boundary is thickly detected. To overcome this problem, the Canny edge detection method is applied to the \( CDM_{n,d}(i,j) \) to extract the rough object shapes. The edge detected version of \( CDM_{n,d}(i,j) \) is called by the edges of significantly different pixels, \( DE_{n,d}(i,j) \). From the value of \( DE_{n,d}(i,j) \), we can grasp the movement of the objects.

2.2 Background registration technique using LL subband

The background registration technique constructs the background from several consequence frames in the image domain, and then extracts a moving object from the difference between the constructed background and the current frame. We consider the LL subband as an image, though it is an approximated version of the image with half-resolution. For the low frequency subband, we apply the background registration technique to build background. The frame difference mask of the LL image, \( FD_{n,LL}(i,j) \), is obtained by thresholding the difference between coefficients in two LL subbands as follows:

\[
FD_{n,LL}(i,j) = \begin{cases} 
1 & \text{if } |w_{n,LL}(i,j) - w_{n-1,LL}(i,j)| > V_{\text{th},FD} \\
0 & \text{otherwise} 
\end{cases}
\]

where \( V_{\text{th},FD} \) is a threshold of \( FD_{n,LL}(i,j) \). If \( FD_{n,LL}(i,j) = 0 \), then the difference between two frames is almost the same. Therefore, it is considered to be a background pixel. According to the frame difference mask of the past several frames, pixels that are not moving for a long time are considered as reliable background. The reliable background, \( BR_{n,LL}(i,j) \) is defined as

\[
BR_{n,LL}(i,j) = \begin{cases} 
BR_{n-1,LL}(i,j) + 1 & \text{if } FD_{n,LL}(i,j) = 0 \\
0 & \text{otherwise} 
\end{cases}
\]

The \( BR_{n,LL}(i,j) \) value is accumulated until \( FD_{n,LL}(i,j) \) holds 0 value. At any time that \( FD_{n,LL}(i,j) \) is changed from 0 to 1, \( BR_{n,LL}(i,j) \) becomes 0. If the value in \( BR_{n,LL}(i,j) \) exceeds a predefined value, denoted by \( L \), then we calculate the background difference mask, \( BD_{n,LL}(i,j) \). It is obtained by taking the difference between the current frame and the background information stored; that is,

\[
BD_{n,LL}(i,j) = \begin{cases} 
1 & \text{if } |BL_{n-1,LL}(i,j) - w_{n,LL}(i,j)| > V_{\text{th},BD} \\
0 & \text{otherwise} 
\end{cases}
\]

where \( BL_{n,LL}(i,j) \) is the pixel value in the current frame that is copied to the corresponding pixel in the \( BR_{n,LL}(i,j) \), and \( V_{\text{th},BD} \) is a given threshold. Now we define the edges of significantly different pixels, \( DE_{n,LL}(i,j) \), as the Canny edge detection version of \( BD_{n,LL}(i,j) \).
In the case of \( BR_{n,LL}(i,j) < L \), we assume that the background is not constructed, so we use a frame difference mask, \( FD_{n,LL}(i,j) \), which is calculated in the first step to extract the edges of significantly different pixels. When \( BR_{n,LL}(i,j) < L \), \( DE_{n,LL}(i,j) \) is obtained by the Canny edge detection of \( FD_{n,LL}(i,j) \). Fig. 2 shows the background registration technique in the wavelet domain.

### 2.3 Merging edge masks and finding moving edge

The obtained four edges of significantly different pixels are 25% of the size of the original frame size. To get the object shape to the size of the original frame, we have to merge the four edges of significantly different pixels. To do so, we apply a union operation to four masks. If we let the merged edge mask be \( DE_{n,F}(i,j) \),

\[
DE_{n,F} = DE_{n,LL} \lor DE_{n,HH} \lor DE_{n,HL} \lor DE_{n,HH}.
\]

We must also consider the method using \( DE_n(i,j) \), \( E_n(i,j) \) and \( E_n(i,j) \)'s neighbor pixel for increasing the number of the extracted \( ME_n(i,j) \). From Fig. 3, we find that moving object edges, \( ME_n(i,j) \) of "Trevor," extracted from the proposed method, are more accurate than the previous method in the wavelet domain [7], which uses only a change detection algorithm.

### 2.4 Extraction of VOPs

We obtained the moving object edge map \( ME_n(i,j) \) detected from \( DE_n(i,j) \), as shown in Fig. 3(b). The VOPs are extracted from the moving object edge map. First, we declare the horizontal candidates if they are inside the first and last edge pixel in each row and the vertical candidates for each column. After finding both horizontal and vertical candidates, the intersection region using logical AND operation is marked as extracted VOPs. For video sequences like "Trevor," which contain only a partial moving object, we declare the image boundary as moving object edge pixels if either horizontal or vertical...
candidates touch the image boundary and conduct the process to find candidates. The false candidates of the extracted VOPs are removed by morphological closing operations. After eliminating false candidates, the noise of the background region is removed and the small uncovered region is covered by a morphological opening operation in post-processing. From Fig.4 we can see the final extracted VOPs.

3. Experiment results

To simulate our proposed method, we performed a Haar wavelet transformation for each image frame. Afterwards, the proposed method was applied to “Trevor,” which is a typical head-and-shoulder video in a 256 X 256 format. From Fig. 3, we find that moving object edges of “Trevor,” extracted from the proposed method, are more accurate than the previous method in the wavelet domain [7]. Since head-and-shoulder type of videos such as “Trevor” have minute movement, the background registration method is of poor quality. That is, the moving object is misunderstood as background. But our method uses both the background registration method and a change detection algorithm to eliminate misleading background. From Fig.5 and Fig.6, our method found more VOPs than the previous method reported in [7]. Our method was also applied to a “hall monitor” in a 288 X 352 format, which is a surveillance type of video with relatively small objects and a complex background. The change detection method shows good results for the head-and-shoulder type of video, but poor results for the surveillance type of sequence. Also we extracted more accurate VOPs as shown in Fig.7 and Fig.8. Fig.9 shows the number of moving object edge for two test videos. From Fig.9, we discovered that our proposed method has more moving edges than those of the method in [7].

---

(a)                                                           (b)                                                      (c)

Fig. 5 VOP extraction for “Trevor” (frame 19): (a) original image; (b) conventional method [7]; (c) proposed method

(a)                                                        (b)                                                   (c)

Fig. 6 VOP extraction for “Trevor”: (a) frame # 23; (b) frame # 69; (c) frame # 97

(a)                                                        (b)                                                        (c)

Fig. 7 VOP extraction for “Hall monitor” (frame # 41): (a) original image; (b) conventional method [7]; (c) proposed method
4. Conclusion

In this paper, we have proposed a moving object segmentation algorithm using a change detection method and background registration technique in a wavelet domain. A low frequency subband has been applied to the background construction; in addition, the change detection method has been made by coefficients in a high frequency subband to overcome the drawbacks of the conventional method. Simulation results show that the proposed method can obtain more accurate moving object edges in both head-and-shoulder and surveillance video sequences.

Acknowledgement

This work was supported by the 2nd-Phase Brain Korea (BK) 21 Program funded by the Ministry of Education, Korea

References


Fig. 8 VOP extraction for “Hall monitor”: (a) frame # 66; (b) frame # 131; (c) frame # 181;

Fig. 9 The number of moving object edge points: (a). “Trevor “; (b). “Hall monitor