References

1 PERONA, P., and MALIK, J.: 'Scale space and edge detection using anisotropic diffusion', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1990, **12**, (7), pp. 629–639

2 ALVAREZ, L., LIONS, P.L., and MOREL, J.M.: 'Image selective smoothing and edge detection by nonlinear diffusion. II', *SIAM J. Numer. Anal.*, 1992, **29**, (3), pp. 845–866

3 IZQUIERDO, E.: 'Stereo matching for enhanced telepresence in 3D-video communications', *IEEE Trans. Circuits Syst. Video Technol.*, 1997, **7**, (4), pp. 629–643

Moving object segmentation in DCT-based compressed video

Salkmann Ji and HyunWook Park

A block-based automatic segmentation algorithm has been developed for detecting and tracking moving objects in DCT-based compressed video. The proposed algorithm segments moving objects with block resolution using the stochastic behaviour of the image blocks in the DCT domain.

Introduction: In general, video signals are compressed to save on the storage and transmission bandwidth of networks. After compression, there are still many situations where further manipulation of such compressed video signals is needed. For example, video transcoding is needed when compressed video signals are transmitted over different networks. Video transcoding is the process of converting one compressed bit-stream into another. Chang [1] developed the manipulation and compositing algorithm of motion-compensated (MC)-DCT compressed video signals. Dogan [2] developed the MPEG-4 and H.263 transcoder for heterogeneous multimedia networks. When a compressed video signal is transmitted to a lower capacity channel, the compressed video signal must be re-quantised to fit the target bit rate of the channel. If the transcoder performs the quantisation of the moving object and background with different quantisation parameters, respectively, the video quality of the moving object and the background can be separately controlled. To discriminate the moving object from the background in a compressed video signal, we have developed a moving object segmentation algorithm for compressed video.

Since DCT and quantisation are block-based operations in MPEG and H.26x video coding, the proposed algorithm segments moving objects with block resolution. The proposed algorithm uses only motion vectors and MC-DCT coefficients, which form a compressed bit-stream, for the segmentation features. Standard video coders usually determine the motion vector using a block matching algorithm (BMA). However, since BMA often fails to determine the true motion vector due to many practical factors (such as affine warping, image noises, and occlusion), the motion vector does not always correspond to the true motion. To overcome the limitation of motion vectors, the proposed algorithm merges spatially similar blocks into a region and investigates the stochastic behaviour of the region.

Proposed block-based segmentation algorithm: Fig. 1 shows the overall block diagram of the proposed algorithm. A region is defined as the set of spatially similar blocks. The spatial similarity of two adjacent blocks is measured using a spatial feature vector, which is defined from the DCT coefficients as follows:

$$\vec{f}^{spatial} \equiv (f_0, f_1, f_2, f_3)$$
$$= \left(v(0,0), \sqrt{\sum_{l=1}^{N-1} v(0,l)^2}, \sqrt{\sum_{k=1}^{N-1} v(k,0)^2}, \sqrt{\sum_{k=1}^{N-1}\sum_{l=1}^{N-1} v(k,l)^2} \right) \tag{1}$$

where $v(k, l)$, $0 \leq k, l \leq N - 1$, are the DCT coefficients of an $N \times N$ image block. In eqn. 1, f_0, f_1, f_2, and f_3 represent the mean, horizontal edge, vertical edge, and diagonal edge of the $N \times N$ image block, respectively. If the difference of two adjacent feature vectors is smaller than a given threshold ($QP \times 4/3$ in this Letter, where QP is the quantisation parameter) determined experimentally, those two blocks become a region.

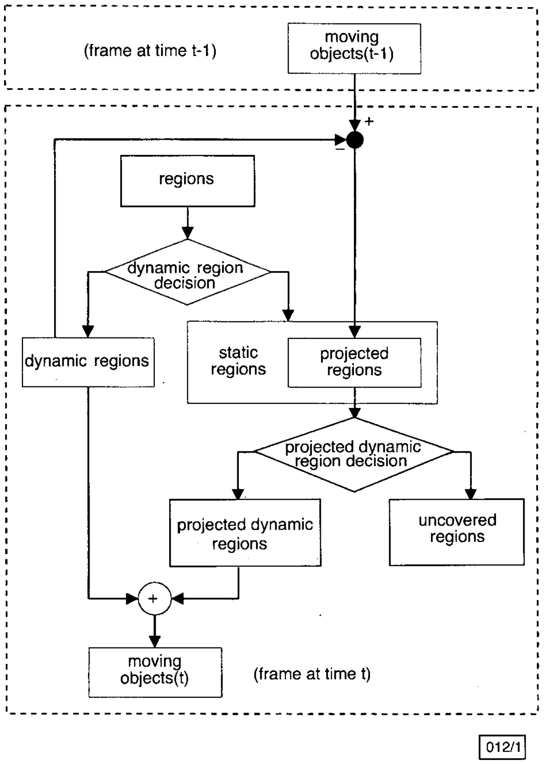


Fig. 1 *Overall block diagram of proposed moving object segmentation algorithm*

In a region with spatial similarity, some blocks are true motion blocks and others are not. Therefore, each block in a region is classified into a true or false motion block using the following decision rule:

$$\begin{aligned} &\text{if } \{e_{t,x,y} \leq d_{t,x,y} \text{ and } e_{t,x,y} \leq d_{t,x-v_x,y-v_y}\} \\ &\quad \mathbf{b}_{t,x,y} \text{ is a true motion block} \\ &\text{else} \\ &\quad \mathbf{b}_{t,x,y} \text{ is a false motion block} \end{aligned} \tag{2}$$

where

$$e_{t,x,y} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \left| b_{t,x,y}(i,j) - b_{t-1,x-v_x,y-v_y}(i,j) \right|^2$$
$$d_{t,x,y} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \left| b_{t,x,y}(i,j) - b_{t-1,x,y}(i,j) \right|^2$$

In eqn. 2, $\mathbf{b}_{t,x,y}$ is an $N \times N$ image block at time t, the upper-left pixel of which is located at (x, y), $b_{t,x,y}(i, j)$ is a pixel value at (i, j) of the block $\mathbf{b}_{t,x,y}$, (v_x, v_y) is the motion vector of $\mathbf{b}_{t,x,y}$, $e_{t,x,y}$ is the motion-compensated error of $\mathbf{b}_{t,x,y}$, and $d_{t,x,y}$ is the inter-frame difference of $\mathbf{b}_{t,x,y}$. The inequalities in eqn. 2 are preserved in the DCT domain according to Parseval's theorem of DCT [3].

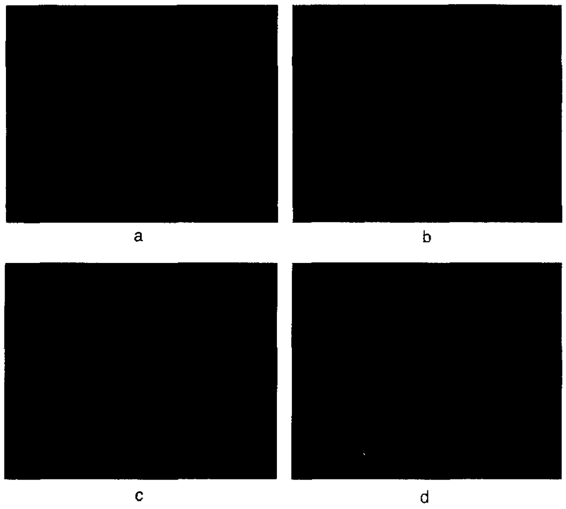
After the blocks have been classified, a dynamic region decision is performed for every region, as shown in Fig. 1. If the number of true motion blocks is larger than that of the false motion blocks in a region, the region becomes a dynamic region. A static region has a smaller number of true motion blocks than of false motion blocks. The dynamic regions represent objects moving in the current frame.

To detect regions that move in the previous frame (time $t - 1$) but stop at the current frame (time t), a projected dynamic region decision is applied to projected regions that are projected from the moving objects in the previous frame onto the static regions in the current frame. The decision rule for determining a true motion block in the projected region is defined as follows:

```
if $\{(\mathbf{b}_{t,x,y} \in$ dynamic region)
    and $(\mathbf{b}_{t,x-v_x,y-v_y} \in$ projected region)$\}$
  $\mathbf{b}_{t,x,y}$ is a false motion block
else
  $\mathbf{b}_{t,x,y}$ is a true motion block                (3)
```

If the number of true motion blocks is larger than the number of false motion blocks in a projected region, the region becomes a projected dynamic region. The uncovered region has a smaller number of true motion blocks than of false motion blocks in the projected region.

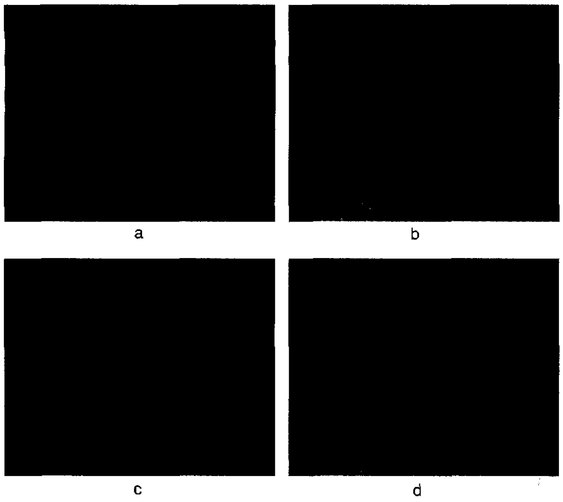
Finally, moving objects consist of dynamic regions and projected dynamic regions.



012/2

Fig. 2 *Blocks with non-zero motion vectors in 'mother and daughter' sequence*

a 3rd frame
b 81st frame
c 180th frame
d 213th frame



012/3

Fig. 3 *Segmented moving object of 'mother and daughter' sequence from proposed segmentation method*

a 3rd frame
b 81st frame
c 180th frame
d 213th frame

Experimental results: Experiments were performed on several compressed bit-streams. Figs. 2 and 3 show blocks with non-zero motion vectors and a segmented moving object, respectively, for the 'mother and daughter' QCIF sequence, which was compressed by an H.263 encoder with a frame rate of 10Hz and a target bit-rate of 24kbit/s. The frame numbers in Figs. 2 and 3 are from the original sequence obtained at a frame rate of 30Hz. Although motion vectors themselves are not appropriate for moving object segmentation as shown in Fig. 2, the proposed segmentation method is able to detect and track moving objects, as shown in Fig. 3.

Discussion: We have developed a block-based moving object segmentation algorithm for compressed video. Since block-based video coders determine motion vectors based on the coding efficiency, motion vectors may give false motion information. However, the proposed algorithm uses the stochastic behaviour of spatially similar blocks to segment moving objects and the segmentation result is successful.

© IEE 2000 *29 August 2000*
Electronics Letters Online No: 20001279
DOI: 10.1049/el:20001279

Salkmann Ji and HyunWook Park (*Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, 373-1 Kusong-dong, Yusong-gu, Taejon 305-701, Korea*)

E-mail: hwpark@athena.kaist.ac.kr

References

1 CHANG, S.F., and MESSERSCHMITT, D.G.: 'Manipulation and compositing of MC-DCT compressed video', *IEEE Trans. Commun.*, 1995, **13**, (1), pp. 1–11

2 DOGAN, S., SADKA, A.H., and KONDOZ, A.M.: 'Efficient MPEG-4/H.263 video transcoder for interoperability of heterogeneous multimedia networks', *Electron. Lett.*, 1999, **35**, (11), pp. 863–864

3 JAIN, A.K.: 'Fundamentals of digital image processing' (Prentice-Hall, 1989)

Cryptanalysis of modified authenticated key agreement protocol

Wei-Chi Ku and Sheng-De Wang

Tseng addressed a weakness within and proposed a modification to the key agreement protocol presented by Seo and Sweeney. The authors show that Tseng's modified protocol is still vulnerable to two simple attacks and describe a new enhancement to the Seo-Sweeney protocol.

Introduction: By using a pre-shared password technique, Seo and Sweeney [1] proposed a simple key agreement protocol which was intended to act as a Diffie-Hellman scheme [2] with user authentication. In the Seo-Sweeney protocol, two parties who have shared a common password can establish a session key by exchanging two messages. The authors also claimed that key validation can be achieved by exchanging two more messages. Later, Tseng [3] addressed a weakness in the key validation steps of the Seo-Sweeney protocol. By replying to the message sent from the honest party, the adversary can fool the honest party into believing a wrong session key. Tseng modified the key validation steps of the Seo-Sweeney protocol and claimed that key validation can be achieved in the modified protocol. In this Letter, we will show that Tseng's modified protocol is still vulnerable to two simple attacks. Additionally, a new enhancement to the Seo-Sweeney protocol will be described.

Tseng's modified protocol: As in the original Diffie-Hellman scheme [2], the system possesses two public values n and g, where n is a large prime and g is a generator with order $n - 1$ in $GF(n)$. Let Alice and Bob denote the two parties who have shared a common password P. The protocol has two phases, the key establishment phase and the key validation phase, and can be described as follows:

Key establishment phase:

(*e.1*) Alice and Bob each compute two integers Q and Q^{-1} mod (n