WAM 12.4

A REAL-TIME PRACTICAL VIDEO SEGMENTATION ALGIRITHM FOR MPEG-4 CAMERA SYSTEMS

Shao-Yi Chien, Yu-Wen Huang, Shyh-Yih Ma, and Liang-Gee Chen
DSP/IC Design Lab, Department of Electrical Engineering
National Taiwan University, Taipei, Taiwan, R.O.C.

ABSTRACT

This paper presents a real-time practical video segmentation algorithm for MPEG-4 video camera systems. The processing speed of 40 QCIF frames per second can be achieved on a personal computer with a 800MHz Pentium III processor. It can be integrated into MPEG-4 based videophone systems and digital cameras.

INTRODUCTION

MPEG-4 standard is designed for interactive multimedia and communication applications. The key function of MPEG-4 is content-based coding, which needs shape information generated with video segmentation systems. Several video segmentation algorithms have been proposed [1-3]; however, most are too complex and too slow for real-time applications. Besides, too many parameters make these algorithms impractical for consumer electronics. A video segmentation algorithm suitable for real-time applications and its implementation are presented in this paper. This algorithm can be integrated into digital camera systems and smart videophones with MPEG-4 encoding systems to bring the content-based coding functions into full play. Besides, it can be implemented efficiently with Intel MMX technology or on other devices with SIMD architectures.

PROPOSED ALGORITHM

The block diagram of this algorithm is shown in Fig. 1. It contains four modes. Basic mode is designed for general situations. Shadow cancellation mode (SC mode) is designed for the situations when influenced by shadow and light changing. Global motion compensation mode (GMC mode) is for slightly moving camera situations, for example, when cameras are held by hands. Finally, adaptive threshold mode (AT mode) is for the situations when environments are dramatically changed.

A. Basic Mode

The frame difference of current frame and previous frame is thresholded with a parameter *Th* to form *Frame Difference Mask (FDM)*. *Background Registration* observes consecutive *FDMs* to extract background information from sequences. The frame difference of current frame and background frame is also thresholded to form *background difference mask (BDM)*. Next, *FDM* and *BDM* are used to form *Initial Object Mask* with *Object Detection*. Finally,

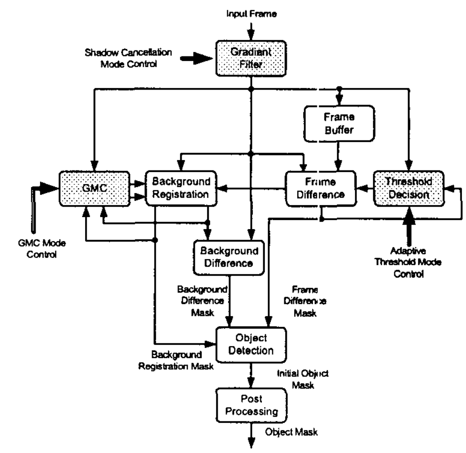

Fig. 1. Block diagram of the proposed algorithm.

post processing, which includes noise region elimination and morphological open-close operations, can improve the *Initial Object Mask* to *Object Mask*.

B. Shadow Cancellation Mode

When the environment is influenced with shadow and light changing, SC mode should be turned on. In SC mode, the *Gradient Filter* in front of the whole system is turned on, which can eliminate the shadow effect and slight light changing effect.

C. Global Motion Compensation Mode

When a camera is held by hands, it will have slight motion, and GMC mode should be turned on. Some feature points are selected according to the gradient value, and then the global motion parameters are estimated from motion vectors of these features points. Then *Background Frame* and *Background Registration Mask* are compensated.

D. Adaptive Threshold Mode

When a camera is first turned on or when environments are dramatically changed, AT mode will be turned on. Since camera noise is modeled as gaussian distribution, a gaussianity test is applied to find the background region roughly. The standard deviation of background region is used to determine the threshold *Th* with the decision curve as shown in Fig. 2, which is derived from digitized gaussian distribution.

0-7803-6622-0/01 $10.00 © 2001 IEEE

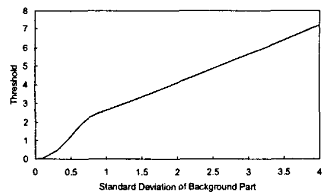

Fig. 2. Threshold decision curve.

IMPLEMENTATION

This algorithm can be further optimized with single-instruction-multiple-data (SIMD) concepts. Since the data-path of state-of-the-art commercial processors is at least 32 bits, if a binary pixel is stored with one bit, 32 pixels can be manipulated at the same time, which can be used to accelerate binary morphology operations [4,5]. Besides, *Frame Difference*, *Background Registration*, and *Object Detection* in Fig. 1 are suitable to be implemented with SIMD instructions, which can be found in almost all commercial microprocessors, such as Pentium, UltraSPARC, MIPS, and ARM.

EXPERIMENTAL RESULTS

The segmentation results are shown in Fig. 3. In Fig. 3(a)(b), the algorithm in basic mode is applied on sequence *Weather*. In Fig. 3(c)(d)(e)(f), the algorithm in SC mode is applied on sequence *Hall Monitor* and *Frank*, which are influenced by shadow effect. As shown in the figures, the shadow effect is eliminated. The experimental results of GMC mode are shown in Fig. 4 with sequence *Shaoyi*. The camera moves in this sequence, which can be observed in Fig. 4(a)(c). Note that AT mode is turned on just at the first frame in these experiments. The run-time analyses of original basic mode, optimized basic mode, SC mode, and GMC mode are shown in Fig. 5. The processing speed of basic mode, SC mode, and GMC mode are 40 QCIF fps, 21 QCIF fps, and 11 QCIF fps respectively.

CONCLUSION

A full-function real-time video segmentation algorithm is proposed in this paper. This algorithm is very practical and can be further optimized for processors with SIMD ability. Experiments show that this algorithm can give good segmentation results in general situations in real-time.

ACKNOWLEGEMENTS

This work is supported in part by SiS Education Foundation.

REFERENCES

[1] MPEG Video Group, "Annex F: Preprocessing and Postprocessing," ISO/IEC JTC1/SC29/WG11 N3056, 2000.

[2] D. Wang, "Unsupervised Video Segmentation Based on Watersheds and Temporal Tracking," *IEEE Tran. On Circuit and Systems for Video Technology*, vol. 8, no. 5, pp.539-546, September 1998.

[3] T. Meier and K. N. Ngan, "Automatic Segmentation of Moving Objects for Video Object Plane Generation," *IEEE Tran. On Circuit and Systems for Video Technology*, vol. 8, no. 5, December 1998.

[4] R. V. D. Boomgaard and R. V. Balen, "Methods for Fast Morphological Image Transforms Using Bitmapped Binary Images," *CVGIP: Graphical Models and Image Processing*, vol. 54, pp. 252-258, May 1992.

[5] S.-Y. Chien, Y.-W. Huang, S.-Y. Ma, and L.-G. Chen, "Efficient video segmentation on SIMD architecture for real-time MPEG-4 systems," in *Proc. of Workshop on Consumer Electronics*, 2000.

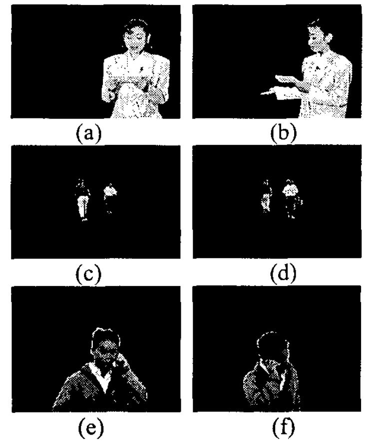

Fig. 3. Segmentation results for sequence *Weather*, *Hall Monitor*, and *Frank*.

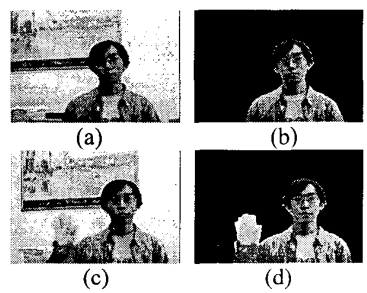

Fig. 4. Segmentation results for sequence *Shaoyi*, when GMC mode is turned on.

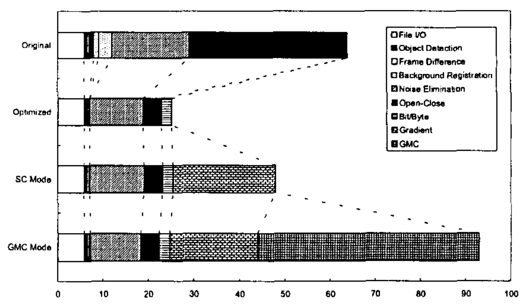

Fig. 5. Run-time analyses