スプライト符号化のための画素単位の動物体抽出
Pixel-based Extraction of Moving Objects for Sprite Coding

渡辺 成光† 秦泉寺 久美†,†† 渡辺 裕†
Shigemitsu WATANABE Kumi JINZENJI Hiroshi WATANABE

†早稲田大学 大学院 国際情報通信研究科 ††NTT サイバースペース研究所
†Graduate School of Global Info. and Tele. Studies, WASEDA Univ. ††NTT Cyber Space Labs.

1. Introduction

MPEG-4 provides an efficient coding tool called ”sprite coding.” In sprite coding, a “sprite” i.e. a unified panoramic background image derived from a sequence having a camera motion, is utilized as a video object plane (VOP) to achieve high coding efficiency. An algorithm that automatically generates the background sprite and the foreground objects has been proposed [1-2]. However, this method approximates the foreground object by using macro blocks. For this reason, unnecessary background regions are extracted as foreground, and the composed image may include visual discrepancy. Thus, a method to extract the foreground objects on a pixel base is necessary. In this paper, we propose an algorithm that automatically extracts the moving objects on a pixel base using background difference and watershed transformation.

2. Object Extraction Algorithm

2.1 Background Difference

First, we calculate the difference image between the original image and the background image that is extracted from the background sprite generated by the conventional method.

2.2 Watershed Transformation

In order to extract the foreground object accurately, image segmentation such that the boundary of the segmented region and the contour of the object coincide must be needed. To satisfy this condition, we use the well-known image segmentation tool, “watershed transformation” [3]. We apply the watershed transformation on the gradient image of the difference image generated in section 2.1. However, watershed transformation tends to over-segment. In order to avoid this problem, we generate a mosaic image by filling the region between the watersheds with the gray value in the original image corresponding to the local minima of the gradient image. The boundaries between two regions of mosaic image are valued with the gray tone difference between these regions. Then, taking the gradient of the mosaic image and deleting the boundaries less than some threshold gives the better watershed image.

2.3 Macro-block mask

The conventional method [1-2] generates a mask image of the foreground object that is approximated by macro-blocks. We use this macro-block mask image to mask the watershed image obtained in section 2.2. By doing so, the unnecessary edges obtained by watershed transformation can be deleted, leaving only the edges around the objects we want to extract. By applying morphological closing process to this masked image, final object mask image is obtained.

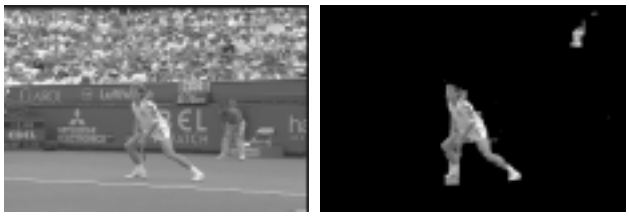

Fig. 1: Original Image (left) and Object Image generated by proposed method (right)

Table 1: Comparison with the conventional method

Frame No	020	040	060	080
MB-based	0.672	0.642	1.384	1.190
Proposed	0.393	0.256	0.475	0.254

3. Result and Evaluation

The original image and the object image generated by our proposed method are shown in fig.1. In order to evaluate the results numerically, we used the video sequence “Stefan” which has the correct object segmentation mask. We evaluated our results by equation (1).

$$e(n)=A_d(n)/A_s(n), \quad (1)$$

where n is the frame number, A_d is the area of regions that is different from the segmentation mask, and A_s is the area of the segmentation mask. Comparison of the proposed method with the conventional macro-block based method is shown in Table 1. It shows that the results given by our proposed method are more than twice as better than that of the conventional method.

4. Conclusion

We proposed an algorithm to extract the moving objects automatically from the video sequence for sprite coding.

References

[1] H. Watanabe, and K. Jinzenji, “Sprite Coding in Object-based Video Coding Standard: MPEG-4,” World Multiconference on SCI 2001 Proc. Vol. XIII, pp.420-425, July 2001.
[2] K. Jinzenji, H. Watanabe, S. Okada, and N. Koboyashi, “Very Low Bit-Rate Video Compression Using MPEG-4 Sprite Coding,” IEICE Transaction, Vol.J84-D-II, No.5, pp.758-768, May 2001.
[3] S. Beucher, “The Watershed Transformation Applied to Image Segmentation”, 10th Pfefferkorn Conf. on Signal and Image Processing in Microscopy and Microanalysis, 16-19 sept. 1991.
[4] S. Sakaida, M. Naemura, and Y. Kanatsugu, “Moving Object Extraction Using Background Difference and Region Growing with Spatio-Temporal Watersheds,” IEICE Transaction, Vol.J84-D-II, No.12, pp.2541-2555, December 2001.