DEFOCUS-BASED IMAGE SEGMENTATION

Cassandra Swain

Vanderbilt University
Box 1824, Station B
Nashville, TN 37235
swain@vuse.vanderbilt.edu

Tsuhan Chen

AT&T Bell Laboratories
101 Crawfords Corner Road, Room 4C528
Holmdel, NJ 07733-3030
tsuhan@research.att.com

ABSTRACT

Foreground and background features are focused (or defocused) differently in an image because corresponding objects are at different depths in the scene. This paper presents a novel approach for segmenting foreground and background in video images based on feature defocus. A modified defocus measurement that distinguishes between high-contrast defocused edges and low-contrast focused edges is presented. Defocus-based segmentation is desirable because defocus techniques are computationally simple. Results indicate that the foreground is easily segmented from moving background. This approach, coupled with motion detection, can segment complex scenes containing both moving background and stationary foreground.

1 INTRODUCTION

Foreground and background segmentation is an important issue in video coding. Typically, the foreground contains important information; whereas, the background does not. So, the background can be transmitted less frequently, which is an advantage in bandwidth constrained transmissions. The problem is how to perform this segmentation simply and accurately.

The primary segmentation techniques are motion-based segmentation [1], intensity-based segmentation [2], and disparity-based segmentation [3]. The motion-based approach segments objects with similar velocities. This approach fails for scenes containing both foreground and background motion, e.g. moving foreground shadows cast onto the background. The intensity-based approach segments images based on intensity contrast and spatial position. This approach fails for textured objects because a single object erroneously segments into multiple objects. The disparity-based approach measures the disparity between stereo images to segment objects. Point correspondence to measure disparity is a complex and error-prone task.

Image defocus is a measure of image sharpness. Researchers use defocus, a function of depth, to determine object distances in monocular systems [4, 5, 6]. Foreground and background image features are defocused differently. This difference can be used to segment and discard background information.

In this paper, a novel approach to foreground and background segmentation based on image defocus is presented. Defocus is measured using a modified approach to distinguish between high-contrast defocused edges and low-contrast focused edges. There are two main advantages in defocus-based segmentation. One, defocus measurements are computationally simple. Two, complex and moving backgrounds can easily be segmented because their defocus is higher than the foreground. This defocus-based approach is combined with motion detection [7] to successfully segment video sequences. Stationary foreground and moving background can be segmented and discarded with combined defocus and motion approaches. Section 2 reviews defocus theory. Section 3 describes the experimental method. Section 4 shows experimental results. Conclusions and future work are presented in Section 5.

2 THEORY

2.1 Image Defocus

Image defocus measures image sharpness. Focusing on an object increases its image sharpness and decreases defocus. Figure 1 shows a simple lens model. Object point N is not in focus. The resulting image is a defocused image n called a *blur circle*. The blur circle size, hence the amount of defocus, is a function of the depth u of point N.

Depth as a function of defocus is [4]

$$u = \frac{sDf}{sD - Df - k\sigma f}, \quad (1)$$

where s is the distance between the lens and the image

0-7803-2431-5/95 $4.00 © 1995 IEEE

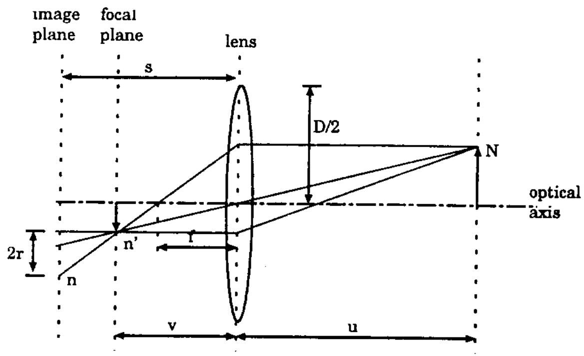


Figure 1: Lens model with blur circle

plane, D is the aperture diameter, f is the focal length, k is the proportionality constant, and σ is the defocus parameter of the blur circle. Parameter σ and defocus vary directly with object depth.

2.2 Defocus Measurement

Image defocus is easily measured from high frequency components, such as image edges. The less blurred an edge, the lower the image defocus. Defocus is measured from edge strength [8]. However, focused edges with low magnitude and defocused edges with high magnitude may give similar defocus measurements. This is incorrect. The defocused edges should have higher defocus measurements.

This error can be corrected by modifying the defocus measurement. The distinguishing feature between focused and defocused edges is edge width. Focused edges have steeper intensity gradients and smaller widths than defocused edges. The defocus measurement can be modified, so that defocus d is

$$d = \frac{\sum_x \sum_y |S(x,y)|^2}{w}, \tag{2}$$

where $S(x,y)$ is the magnitude of Sobel edge detection on image $g(x,y)$ and w is the edge width in $g(x,y)$.

3 METHODOLOGY

An defocus-based image segmentation approach has been developed. Foreground and background features are segmented and background features discarded. The results of this segmentation are combined with motion detection to extract the foreground in an image sequence. This work assumes that the foreground is focused and the background is defocused and that the image contains high frequency features.

Figure 2 shows the five step segmentation algorithm. First, an image is acquired. Next, defocus is measured and the image segmented based on defocus. The foreground is retained and the background discarded. Motion is also detected. The moving features are retained and the stationary ones discarded. The third step is the filling operation. Defocus and motion are measured at object edges. However, object interiors must be segmented also. Multiple-pass filling is performed to include the object interior, preserving object shape. Next, the segmented defocus and motion results are intersected to get the common region. This region is the foreground. The segmented foreground is then outlined on the original image.

4 EXPERIMENTAL RESULTS

4.1 Segmentation

Figure 3 shows the segmentation results for an image sequence. Note the defocus differences between the human subject in the foreground and the objects in the background. Also, some foreground and background features were in motion.

The motion algorithm detected all moving objects, the subject in the foreground and the hand in the background. The defocus algorithm segmented the focused object, the subject in the foreground. The common region, the subject in the foreground, was the final result.

This test sequence does not contain stationary foreground features. However, any stationary foreground would be segmented by the defocus algorithm as foreground, but discarded by the motion algorithm. The final result would be segmentation of the moving foreground, which is desired.

4.2 Defocus Measurement

Figure 4 shows segmented features using the old and the modified defocus measurements. The moving hand in the background in 4(a) had high contrast. Though defocused, it had a high defocus measurement. Using the old defocus measurement, the hand was segmented as foreground, as seen in the upper left corner of 4(b). This error was corrected by using the modified measurement, as the segmentation in 4(c) shows. Since the defocused background had wider edges than the focused foreground, the modified measurement, Equation (2), gave the correct defocus measurement.

5 CONCLUSIONS AND FUTURE WORK

A novel approach to foreground and background segmentation has been presented. Image defocus, a function of depth, is a basis for segmentation. Defocus-based segmentation is advantageous for two reasons. One, it is computationally simple. Two, scenes con-

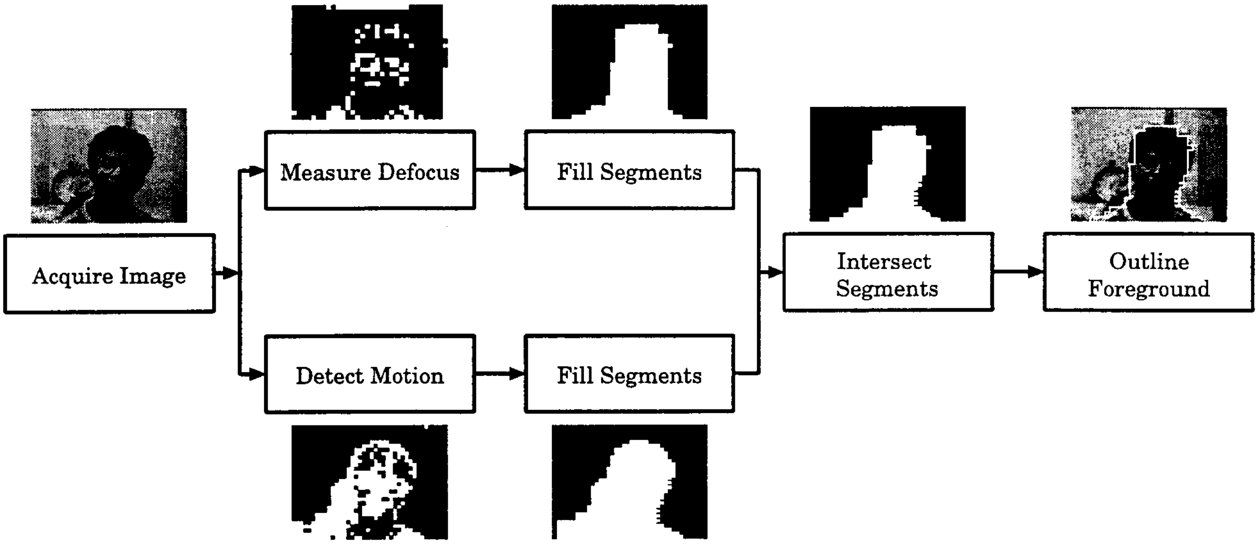
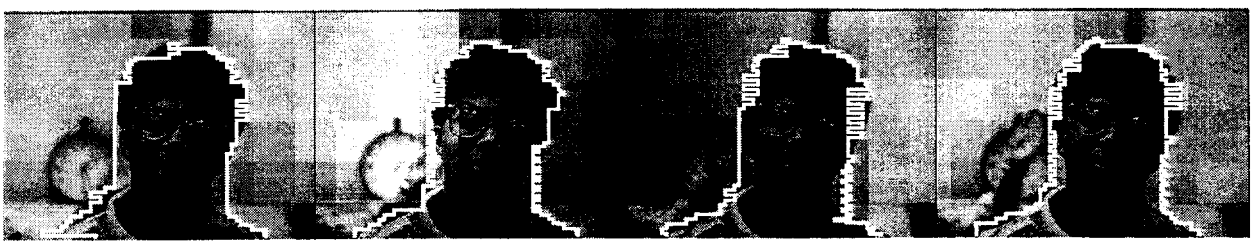


Figure 2: Foreground and background segmentation method



(a) (b) (c) (d)

Figure 3: Segmentation results (a) Frame 1. (b) Frame 2. (c) Frame 3. (d) Frame 4

(a) (b) (c)

Figure 4: Results of defocus measurement modificiation

taining complex or moving backgrounds can be segmented easily.

Defocus-based segmentation combined with motion detection segments scenes containing multiple foreground objects.

Future research will address the following issues:

- Include color and intensity-based segmentation with the defocus-based approach. As a result, the filling operation will not be necessary.
- Develop alternate methods for combining defocus and motion detection results. In this approach, defocus and motion segmentation were performed separately and then combined to get the intersecting region. An alternate method would combine defocus and motion into a single segmentation step.
- Use temporal correlation between adjacent frames in assisting defocus-based segmentation.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Hans Peter Graf, AT&T Bell Laboratories, and Dr. Richard Alan Peters II, Vanderbilt University, for their technical assistance.

References

[1] S. Peleg and H. Rom, "Motion Based Segmentation", *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 109-113, 1990.

[2] R. Haralick and L. Shapiro, "Image Segmentation Techniques", *Computer Graphics, Vision, and Image Processing*, Vol. 2, pp. 100-132, 1985.

[3] M. Waldowski, "A New Segmentation Algorithm for Videophone Applications based on Stereo Image Pairs", *IEEE Trans. Communications*, Vol. 39, No. 12, pp. 1856-1868, 1991.

[4] A. Pentland, "A New Sense of Depth of Field", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 9, No. 4, pp. 523-531, Jul 1993.

[5] M. Subbarao and G. Surya, "Depth from Defocus: A Spatial Domain Approach", *Technical Report No. 92.12.03*, Computer Vision Laboratory, Electrical Engineering Department, SUNY, Stony Brook, NY.

[6] C. Swain, M. Bishay, A. Peters, and K. Kawamura, "Accuracy Improvement of Depth from Defocus using Fuzzy Logic", *Technical Report No. CIS-94-02*, Center for Intelligent Systems, Vanderbilt University, Nashville, TN 37235, May 1994.

[7] T. Ebrahimi, "A New Technique for Motion Field Segmentation and Coding for Very Low Bitrate Video Coding Applications", *Proc. IEEE Int. Conf. on Image Processing*, pp. 433-437, 1994.

[8] E. Krotkov, "Focusing", *Int. Journal Computer Vision*, Vol. 1, pp. 223-237, 1987.