視野角の異なる二台のカメラを用いた移動物体認識*

1G－2　　黄　英傑 **　　土肥　浩 **　　石塚　満**

東京大学工学部電子情報工学科**

1. Introduction

One of the features of human visual system is the non-uniformity of visual acuity across the visual field. This allows humans to achieve high resolution vision where needed without the costs of uniformity high resolution. To fit the both needs of wide viewing-angle and high resolution, We construct a vision system with dual viewing angles. The wide viewing angle camera is mounted on a fixed position to offers a wide viewing field but at low resolutions. Another camera with narrow viewing angle, thus high resolution, is mounted on a pan/tilt platform and then can be turned to maintain a fixation on a moving object.

2. Gaze control system

The gaze control system is used to drive the narrow viewing angle camera to keep the image of object at the center of its viewing field. The configuration of the two cameras is shown in Fig. 1. The relation between the pan-tilt angles and the position of object must be known. This relation, is also called inverse kinematic equations, can be conducted from Fig. 1 by using stereo triangular as follow:

$$X = L\tan\beta$$
$$Y = -L_p\sin\varphi \quad \text{, in which} \quad L = \frac{l + r\sec\theta}{\tan\theta - \tan\beta}$$
$$Z = L_p\cos\varphi \qquad L_p = L\sec\theta - r\tan\theta$$

where θ, ϕ are the angles of pan and tilt respectively, L is the distance from the object to the base line of two cameras, r is the rotation rasius of pointable camera.

3. Motion tracking with Wide viewing angle camera

For tracking to be as general as possible, it should be able to follow a moving object whose identity is not known, i.e., not require a object recognition. The modified motion energy (mme) [1] is one of these methods that are suited for this purpose. By calculating the mme between the current image and the background image, the existence of the object can be detected and can be extract from background at the same time. Fig. 2 shows the results of tacking a person who entered the viewing field of wide viewing angle camera. The head of person can be further extracted by using the geometry knowledge of human body. According to the position information calculated from wide viewing field image, the pointable camera is turned to follow the object to get its image with high resolution as shown in Fig. 2(c).

4. Mouth open/close detection with narrow viewing camera

Face recognition has been found in several applications range from static matching of controlled photographs such as mugshot matching and credit card verification to surveillance video image [2]. It is generally required to segment the face image from a background prior to the recognition procedure. The resolution in Fig. 2(c) is suitable for these works. As an example, we present a method to detect the open/close of mouth on an image sequence of face took from pointable camera.

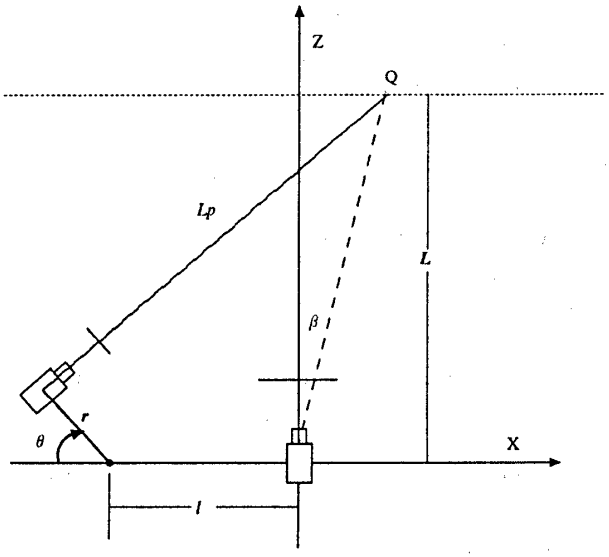

Fig.1 The configuration of the two cameras

Since the shape of the region between lips meaningfully express the movement of lips and this region is easier to be extracted than the extraction of lips directly form face, the open/close of mouth can be

* Moving object recognition with dual viewing angles cameras

** Ying-jieh huang, Hiroshi Dohi and Mitsuru Ishizuka
University of Tokyo, Faculty of Engineering,7-3-1, Hongo, Bunkyo-ku, Tokyo 113, JAPAN

detected by measuring the shape variation of the region between lips.

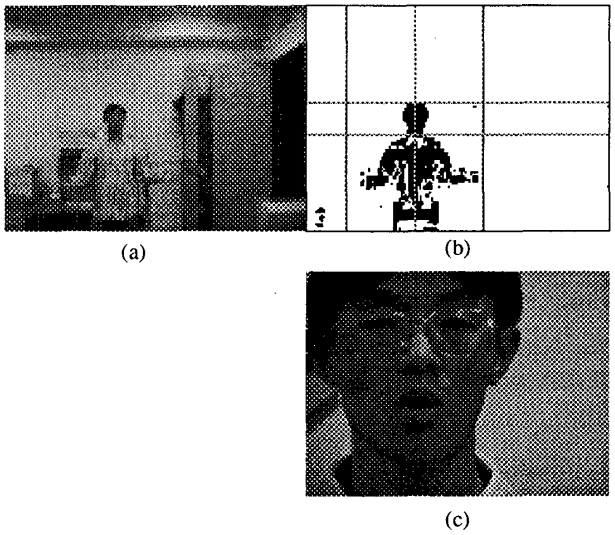

Fig. 2 Motion detection and segmentation in wide viewing field

To extract the region between lips, the face region is segmented at first. When use image format of YIQ, the face region and lip region express peak value in I and Q component respectively [3], and Y component will be use to calculate the threshold value for extract the region between lips. Fig. 3 shows the final results including some intermediate results.

The measuring on the size, length and width of region between lips can be used to determine the open/close of mouth. In Fig. 4, the difference of region between lips between opened mouth and closed mouth in size, length and width is shown.

5. Conclusion

An vision system with dual viewing angles is proposed to implement the human like fovea-peripheral visual acuity. With two cameras of different viewing angle, the motion and the feature of the object can be took at the same time. There are many applications in human computer interface can be applied such as in speech recognition system, hand gesture or finger tips recognition, human gaze detection.

Reference

[1] Y. Huang, H. Dohi and M Ishizuka,"A Realtime Visual Tracking System with Two Cameras for Feature Recognition of Moving Human Face," 4th IEEE International Workshop on Robot and Human Communication(RO-MAN'95), July 1995.

[2] R. Chellapa, S. Sirohey, C. L. Wilson and C. S. Barnes, "Human and Machine Recognition of Faces: A Survey," CAR-TR-731, 1994.

[3] S. Akamatsu, T. Sasaki, H. Fukamachi and Y. Suenaga, "Automatic Extraction of Target Images for Face Identification Using the Sub-Space Classification Method," IEICE Trans. Information and System, Vol. E76-D, No. 10, Oct. 1993.

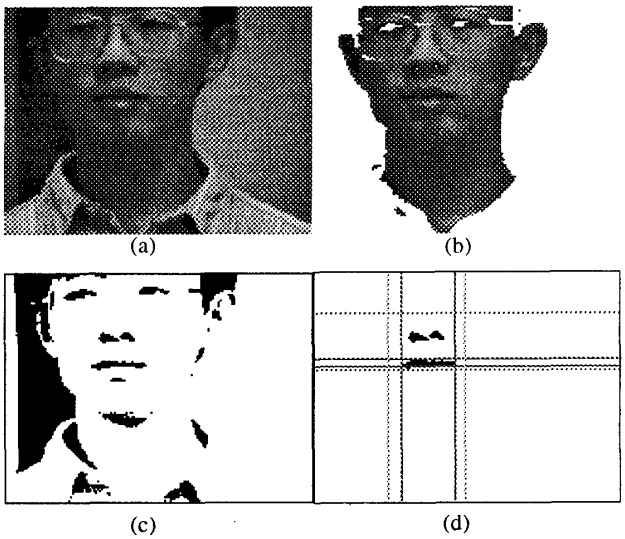

Fig. 3 The region between lips segmentation result

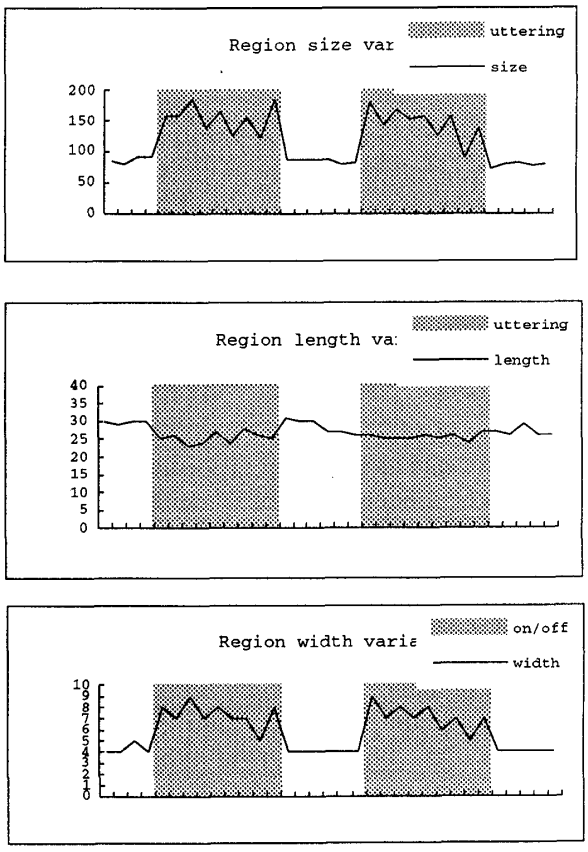

Fig.4 The region between lips of mouth