TRACKING A MOVING OBJECT
WITH A STEREO CAMERA HEAD

S. Rougeaux†, O. N. Kita*, Y. Kuniyoshi* and S. Sakane*

Autonomous Systems Section
Electrotechnical Laboratory, Tsukuba, Ibaraki 305 JAPAN* Instut d’Informatique d’Entreprise, 18 Allee Jean Rostand, 91000 Evry, FRANCE†
email: rougeaux@cnam.cnam.fr, [kitao.kuniyosh,sakane]@etl.go.jp

Abstract

We present a stereo active vision system which performs tracking tasks on smoothly moving objects in complex backgrounds. Dynamic control of the camera vergence adapt the horopter geometry to the target position and allows to pick it up easily on the basis of stereoscopic disparity features. We introduce a novel vergence control strategy based on the computation of “virtual horopters” to track a target movement generating rapid changes of disparity. We then demonstrate the efficiency of the system with experimental results on a two degrees of freedom binocular head performing gaze fixation on a smoothly moving target translating and rotating in a complex surrounding.

1 Introduction

The importance of eye movement to biological visual systems is obvious. In contrast, controlled camera movement have played a small role in computer vision research, but are becoming increasingly recognized as important capabilities in robotic visual perception (see, in particular the work of Ballard [1] and Christensen [2]). In fact, active motion of camera may provide many advantages: Since tracking involves that the visual target remains near the center of the image, it allows the use of localized visual processing and stereo algorithms that accept only a limited range of disparity. Moreover, since the eyes follow the target, the target image tends to have slow motion across the retina. On contrary, surrounding distractors move rapidly across the retina and suffer from motion blur. As a result, the signal of the target is emphasized over the background.

In binocular systems, whose cameras have their optic axis in the same plane, gaze holding is the process of adjusting pan angles so that both eyes are looking at the same world point. Gaze control may be broken into two different tasks: gaze stabilization, which involves maintaining fixation on a moving (or motionless) visual target from a moving (or motionless) gaze platform and gaze shift, which represents the ability of the vision module to transfer fixation from one visual target to another. We describe in this paper a method for holding gaze on a smoothly moving objects with a binocular head whose vergence is real-time controlled, and show experimental results.

2 Tracking within a horopter

Our tracking algorithm is based on stereoscopic disparity features; we assume that the target is initially at the fixation point. Since during successful tracking the target is always near the gaze point, its images keep mostly zero disparity.

![Figure 1: View of binocular fixation](image)

By definition, any point of zero disparity projects onto left and right image points with identical coordinates. The set of such points (the horopter, shown in figure 1) is located on a circle, also called the Vieth-Müller circle, passing through the two nodal points of the cameras and the gaze point. Thus, objects which stay in this horopter can be easily picked up from many other objects.
by suppressing features of non-zero disparity. This well known method has been in particular implemented for the Rochester robot head [3]. The principle of the zero disparity filter (ZDF) is simple: we first extract vertical edges from both left and right images by Sobel operation. Then, the stereo edge images are compared in corresponding pixel location. As a consequence, only edges of objects lying in the horopter may remain on the matching output. Considering that only one object is located on the common space of both cameras field of view and the horopter, we just have to compute the center of the gravity mass of the ZDF output to obtain a rough measurement of the target location.

3 The virtual horopter

Unfortunately, the strategy of zero disparity filtering is efficient for object moving exclusively along the horopter. Recently, Coombs and Olson [4] developed a method based on cepstral filtering [3] for measuring the disparity error, but experiments showed that this process requires the use of powerful hardware in order to keep a real-time feature. We propose a novel approach for the localization of object moving across the horopter: to cope with vergence error, we extend ZDF with a simple algorithm based on the computation of virtual horopters. This virtual horopter is the horopter generated by shifting horizontally the right image by a certain amount of pixel. Low shift values are equivalent to small virtual rotations of the right camera. Doing so, we obtain a new virtual gaze point and a new virtual horopter.

4 Head design

Our gaze platform has two degrees of freedom corresponding to the two rotations of the cameras around the vertical axis. It is equipped with DC motors and potentiometers for pan angle range measurements.

The stereo image acquisition is performed by two CCD color cameras. However, due to the limitations of our image processing system, we only use gray scale information. This platform is a prototype designed to be easily manipulated and mounted on a robot’s arm effector or on a small autonomous mobile robot: the size of the global system, including cameras, is about 250 mm width, 150 mm height, 170 mm depth and weight is about 1 Kg.

5 Implementation

The tracking system is managed by MaxTD, a workstation composed by a LynxOS kernel and Maxvideo 200, a real-time image processing hardware. Stereo image acquisition is performed by Digicolor using red and green components of the video signal. Figure 4 clearly shows the flow of data in the system between the different devices.
The workstation is also connected through a serial port to a controller board which holds a classic PD control loop for managing cameras rotation around pan axis.

Table: Gaze platform

<table>
<thead>
<tr>
<th>Cameras</th>
<th>Potentiometers</th>
<th>Amplifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motors</td>
<td>VME</td>
<td>Controller</td>
</tr>
<tr>
<td></td>
<td>Digicolor</td>
<td>V25 (8086) MHz</td>
</tr>
<tr>
<td></td>
<td>Maxvideo 200</td>
<td>RS232</td>
</tr>
<tr>
<td></td>
<td>LynxOS (68040) 25 MHz</td>
<td></td>
</tr>
</tbody>
</table>

**Datacube MaxTD**

![Diagram of gaze platform](image)

**Figure 4: Overview of the stereo tracking module**

6 Performance

Central 230x130 portions of 512x480 gray scale images digitized on 8 bit per pixel are processed by the ZDF filter. It takes 1/45 sec for the whole process (stereo acquisition, edges extraction and the 3 ZDF matching), then there is still some rest for other processing to achieve frame rate (30 Hz). We experienced that this module was able to track object moving around \(50^\circ s^{-1}\) along the horizon and around \(15^\circ s^{-1}\) across it. Figure 5 shows both camera pan angle values during a smooth moving target tracking task.

This system also roughly estimates the target position in the 3D world environment (±6 cm in depth for a target moving at 1 meter from the baseline of the gaze platform). The inaccuracy is essentially caused by the noisy output of the potentiometers reporting cameras pan angle rotation: it has around ±0.4° error.

7 Conclusion

This work describes a simple method based on zero disparity filtering for tracking a smoothly moving object in complex surroundings. We have shown how a simple method based on the computation of a virtual horizon may help to determine the vergence error and thus permits the tracking of object moving either along or across the real horizon. We already implemented this stereo tracking module in a real-time application on the Maxvideo 200 hardware running at 30 Hz.

However, one major drawback of this algorithm is that blurring raises the chance that edges match in the ZDF output by accident even if those object aren’t in the horizon. Further improvements to avoid such wrong matching are necessary to make tracking in complex backgrounds more robust. Moreover, in the current situation, the system requires that the target object is at the fixation point for each initialization of the tracking task. To avoid this drawback, we already started to develop a target acquisition algorithm based on optical flow \[6\] in order to acquire any moving object passing through the field of view of both cameras. In cooperation with the above method, it could constitute a surveillance loop \[7\] able to cope with gaze shift and gaze stabilization tasks. In the near future, we plan to develop a new version of this gaze platform, which will allow pan and tilt rotation for both cameras and a neck rotation for the whole head system.

![Graph of motor response](image)

**Figure 5: Motor response while tracking a target moving on a record player rotating at 0.4 Hz, with a radius of 30 cm and a distance from the gaze platform baseline of 1 meter. Error means the angle difference between the optimal (target) position and the real positions of the cameras.**

References


Figure 6: View of the monitoring display. The rectangles drawn on the edges images show the input matching areas for the ZDF processing. You can notice that the right image has been shifted to the left in order to correct the vergence error. The lower left corner of the display shows the estimated target position and both real (continuous line) and virtual (dotted line) horopters. This figure clearly shows that the computed virtual horopter is well adapted to the target position.

Figure 7: Three different frames of the left camera during a tracking sequence.