Proceedings of ICAR 2003
The 11th International Conference on Advanced Robotics
Coimbra, Portugal, June 30 - July 3, 2003

Design of a low-power motion tracking system

Kevin Nickels, Chuck Divin, Jason Frederick, Luke Powell, Christina Soontornvat
knickels@trinity.edu
Dept. of Engineering Science
Trinity University
San Antonio, TX 78212

Jeffrey Graham
jgraham@lincom-asg.com
NASA/JSC Robotics and Automation Group
Titan-Lincom
Houston, TX 77058

Abstract

Mobile robots require fast, low-power motion tracking systems. Popular systems require much power to move the mass of two cameras on a pan-tilt-verge head, and thus are not suited well for battery-powered mobile robots. In this paper, we present a new system for motion tracking that utilizes a mirror-based optical system to produce a stereo image on one camera. Processing one image solves synchronization challenges common in stereo systems and requires fewer computing resources than processing two images. We have developed a system that uses about $1/50th$ *of the power of a popular system with similar performance specifications.*

1 Introduction

The aim of this project is to develop a stereo vision system (shown in Figure 1) capable of motion tracking at standard frame rates while minimizing power consumption. The speed of tracking can take on two meanings in a tracking system: the speed of the overall computation and saccade cycle, or the saccade velocity and acceleration. We emphasize the second meaning, as our current system utilizes a standard analog-output CCD camera and is therefore limited to 30 frames per second.

Stereo vision systems are important for motion tracking applications, because two images are used to compute the location of an object in the world and track the object. Many current motion tracking systems are comprised of two spatially separated cameras mounted on a pan-tilt (and sometimes verge) head, two frame grabbers to digitize the images, and a computer to process the images. A prototypical example is TracLabs' Biclops [7] shown in Figure 2.

These systems pose several challenges for the developer. First, many stereo algorithms assume that the two images are taken at exactly the same point in time. While this can be achieved with modern cameras, it is cumbersome to implement. Secondly, while small and light CCD cameras are available, they are still expensive. The movement of two more typical cameras, lenses, and the supporting framework at the desired velocities consumes a fair amount of power, which is always at a premium in mobile applications.

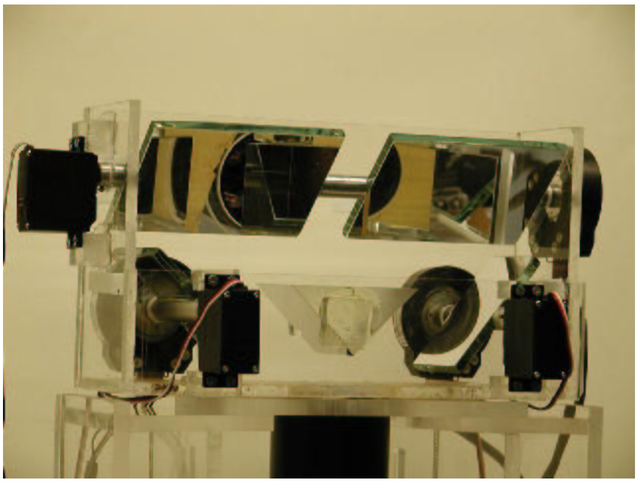

Figure 1: *Stereo Vision Motion Tracking System*

We have developed the stand-alone lightweight stereo imaging system shown in Figure 1, capable of motion tracking. The system has low power requirements, but maintains the view acceleration and velocity specifications of comparable systems.

To reduce power requirements, simplify the image processing, and solve synchronization challenges, one camera has been used to obtain a stereo image instead of two. The use of one camera requires an optical system to project two spatially separated images onto one image sensor.

Several solutions to the single lens stereo problem have been proposed. One approach involves using a biprism to refract the images and view two separate images while looking through the biprism [2]. While the biprism solution produces a stereo image, the field of view is constrained by the phys-

Figure 2: *The Biclops Stereo Head*

ical size of the biprism. Additionally, the camera and the biprism would have to move together and remain aligned throughout the tracking process, reducing the improvements in power needs from the use of one camera.

Goshtasby and Gruver present a method to change a stereo vision system's field of view by using hinged mirrors [1]. Their system effectively produces a stereo image on one camera; however, the system requires that the camera be mounted above and in front of the mirrors. The camera is difficult to mount and the view could be obstructed with the camera in front of the mirrors.

Teoh and Zhang present a system that utilizes a rotating center mirror to project a left and right image to one camera [6]. This design requires that the scene remain the same between taking pictures, because the center mirror must be rotated to achieve the stereo image. This system therefore cannot be used to track a quickly moving object.

Nene and Nayar have proposed four stereo systems using planar, hyperbolic, ellipsoidal, and parabolic mirrors [4]. Their systems need gimbal joints to mount the mirrors, and they require complex transformation algorithms, particularly for the non-planar mirrors.

Pentax Corporation and Pieter Zanen have independently invented 3-D stereo imaging systems that focus a left and a right image onto one camera chip by using two angled mirrors and a wedge mirror [8]. The Pentax system uses fixed angled mirrors and is designed to be attached to a standard film camera. Zanen's system incorporates two fixed mirrors that are coupled with a linear gear drive. Zanen's system, therefore, only has verge motion.

Our system uses four angled mirrors. Two of these mirrors rotate independently and can produce pan and verge motion. The other two mirrors are coupled on a rotating axis, and produce tilt vision.

2 Hardware Description

This section describes our single lens stereo motion tracking system. Our system capitalizes on the advantages of single lens stereo via low inertia optics to track moving objects. Figure 3 presents our approach to the problem.

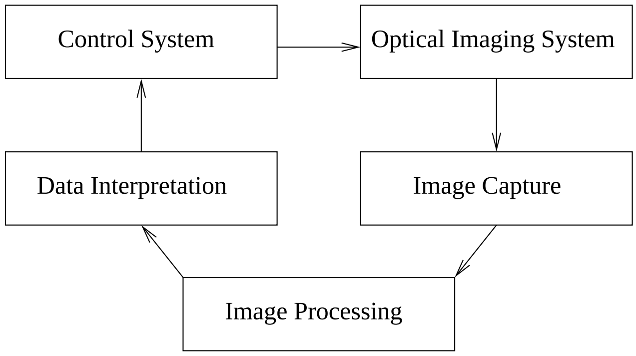

Figure 3: *Overview of System*

A control system orients an optical imaging system to frame a target object. A camera captures this image and transmits it to an image processing module that searches for motion. The processed image data is sent to a data interpretation system that determines a new set point for the control system, the control system is notified of the new set point, and the cycle repeats.

2.1 Optical System

The optical system incorporates four planar mirrors and one wedge mirror as displayed in Figure 4. The two upper mirrors rotate about the same axis. This axis controls the tilt motion. Due to manufacturing errors, these mirrors are not at precisely the same angle (see Figure 5,) but this does not degrade system performance significantly. The two lower mirrors rotate independently on parallel axes. These axes control the system's pan motion. By limiting the motion of each mirror to one degree of freedom, they are easier to control, and the mounting joints are lighter than a system of mirrors with multiple degrees of freedom.

The mirrors are attached with an industrial epoxy to aluminum shafts, which have been milled in the center to provide more surface area to attach the mirror and to place the center of rotation close to the mirror's center of mass, which reduces inertia. The aluminum shafts are directly attached to the servo motors.

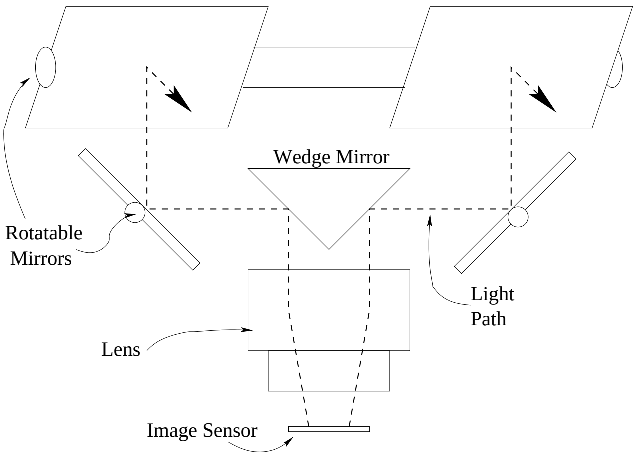

Figure 4: *Mirror Based Optical System*

Figure 5: *Sample Camera View*

These mirrors are placed above a Navitar Zoom 7000 lens and a Pulnix CCD TMC-7DSP camera. The camera interfaces with a Hauppage WinTV framegrabber and a general purpose PC for image processing and control signal generation.

Figure 1 shows the prototype design, from the lens up. The camera and lens are attached vertically below the stereo vision module so that the lens looks directly upward at the wedge mirror.

2.2 Control System

Each axis of rotation is controlled by a dedicated servo motor. Servo motors were chosen over stepper motors primarily because they require less power. Three U.S. Digital optical encoders with an effective 4000 counts-per-revolution are attached to the shafts to monitor their rotation. These encoders require a single 5V supply.

An MIT Handyboard [3] controls the motors and receives the encoder response. The Handyboard is based on a Motorola 6811 microprocessor. It communicates with the main computer via an RS-232 serial connection.

3 System Performance

This section describes the experimental tests that were used to characterize the system performance. During development and testing, the several issues were uncovered that had to be addressed. These issues are discussed in this section. The tracking results follow in Section 3.5.

3.1 Positioning Accuracy and Precision

Initial tests revealed hysteresis in the control of the mirrors, as the three shafts and mirrors did not return to the same start position after a forward and reverse rotation. The hysteresis is attributed to the slip between the gears in the servo motor. The average difference between the start and end positions for each mirror was 1.4°. With the hysteresis, the system could not be controlled to the desired 10 arc-min accuracy resolution. This system is designed to use direct drive of the mirrors, without the gear train often used to increase positioning accuracy and precision.

This hysteresis problem was reduced with a start-up calibration procedure in software. The calibration procedure determines the amount of slip at start-up, and then corrects future measurements based on the initial slip. This calibration procedure takes about 15 seconds and reduces the control hysteresis by an average of 88% to 0.17° for tilt motion and an average of 82% to .25° for pan motion. Figure 6 shows the positioning accuracy of the tilt axis before and after software calibration.

3.2 Image Interference Region

A small vertical band in the center of the image (about 44 pixels horizontally, or around 1.5°) is unusable because of an interference region. This interference region results in a superposition of the two stereo images due to light diffraction around the edges of the wedge mirror.

This interference region effected approximately 7% of the image, but did not significantly affect our tracking ability within the visible region. The interference region can be seen as the blurred area in the middle of the sample image in Figure 5.

3.3 Field of View and Range of Motion

The instantaneous vertical field of view (FOV) of the system is estimated at $\pm 7.5°$. The horizontal FOV is estimated at $\pm 4.1°$ for the left image and $\pm 6°$ for the

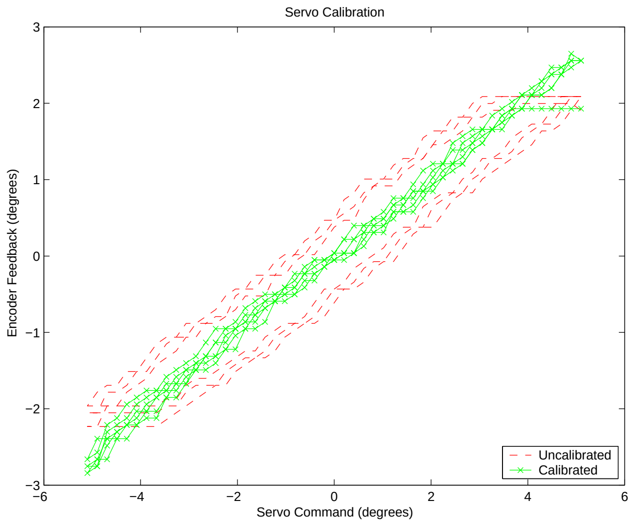

Figure 6: *Positioning Accuracy of mirrors*

right image. As can be seen in Figure 5, the wedge mirror is not perfectly centered on the camera lens, leading to this asymmetry.

The primary reason for the narrow FOV is the placement of the moving mirrors in front of the standard camera optics. A secondary explanation for the small FOV is the use of a single CCD array to capture two images, leading to some inherent restrictions. We believe that the limited FOV is more than compensated for by the simplicity given by a single camera and the power savings realized by the lightweight moving optics of the system.

The total visible pan angle range (the angle between the leftmost point visible at any configuration and the rightmost point visible at any configuration) is estimated at $+38.5°/-34.3°$ for the left image and $+19°/-32°$ for the right image. The total visible tilt angle range is estimated to be $+32.5°/-31°$.

3.4 Motion Specifications

Table 1 presents our system specifications. These specifications meet or exceed the specifications of currently available pan-tilt units, with the exception of the range of motion.

3.5 Tracking

This section describes the tracking performance of the stereo system in a closed loop with image processing. The tracking results were obtained by tracking with the left view of the image only. One view suffices because these tests were completed on a two dimensional background.

Our system successfully tracked a black circle on a white background. Color tracking on general backgrounds has also been implemented. The object was placed at a distance of eight feet from the camera. Figure 7 shows the results from tracking the black circle as it moved in a square pattern. The figure shows the path of the circle as measured by the computer and the independent ground truth as determined by hand measurement. The system tracked the object within a maximum deviation from the path of 2 cm.

Table 1: *Specifications*

Category	Pan	Tilt
Range of Motion	$\approx \pm 30°$	$\approx \pm 31.75°$
Inst. Field of View	$\approx \pm 5°$	$\approx \pm 7.5°$
Max Velocities	$290°$/sec	$302°$/sec
Max Accelerations	$9180°/\text{sec}^2$	$10300°/\text{sec}^2$
Resolution	10 arc-min	10 arc-min
Power	≤ 3 W	

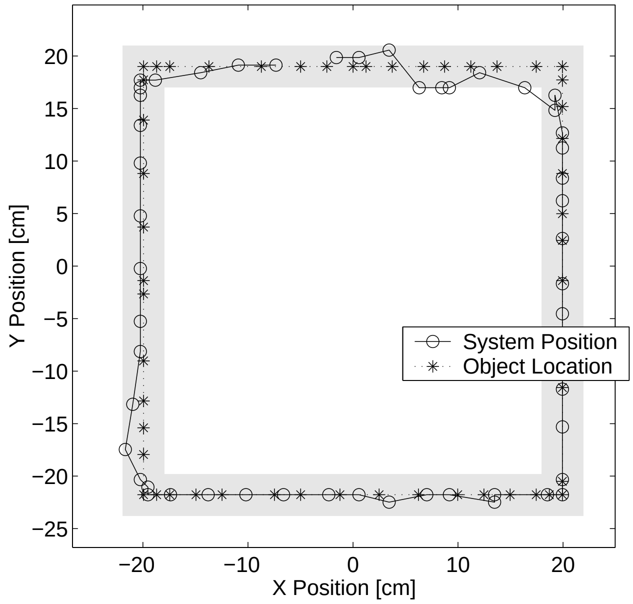

Figure 7: *Motion Tracking.*

The tracking cycle is currently restricted to approximately 20 frames per second. The primary limitation on the speed of this cycle is the MIT Handyboard. This microcomputer (68HC11) supports several baud rates in excess of 9600 baud, but none are compatible with what are now standard PC baud rates. If the baud rate were increased, 30 frames per second could be achieved with the current cameras.

3.6 Performance Comparison

Table 2 presents the maximum angular velocities and acceleration of our system and compares that to two other commercially available systems. These two systems have verge capability. Our system implements independent pan on each camera, so no explicit verge specifications are given.

The power calculations include the motor and logic power requirements only. If cameras were included, our system would accrue more marginal benefit, as we require one camera in contrast to the typical two. The given velocities and accelerations are *unloaded*. That is, they do not include the weight of cameras or lenses. In the case of our system, there is no change for loaded performance since the camera is stationary.

Our system has a response for both pan and tilt motion that exceeds the Biclops specifications and nears or exceeds the Zebra specifications. Our system's power requirements, however, are less than 1/6th of the Biclops power and less than 1/50th of the Zebra power. The numbers in Table 2 are the maximum velocity and acceleration for visual rotation.

Table 2: *Comparison of Specifications*

	Our System	Biclops[1]	Zebra[2]
Range of motion			
Pan	$\approx \pm 30°$	$\pm 165°$	cont.
Tilt	$\approx \pm 31.75°$	$\pm 60°$	90°
Maximum velocities			
Pan	580°/sec	120°/sec	360°/sec
Tilt	600°/sec	120°/sec	270°/sec
Maximum accelerations			
Pan	9180°/sec^2	300°/sec^2	8264°/sec^2
Tilt	10300°/sec^2	300°/sec^2	8264°/sec^2
Power	$\leq$ 3W	$\leq$ 20.25W	$\leq$ 151.5W
Resolution	10 arc-min	n/a	n/a

4 Conclusions and Future Work

This system effectively tracks motion. However, the system is constrained by a limited field of view and range of motion. Future work will address communication speed and field of view considerations. A more modern controller might be used to increase the (serial) data rate of control, effectively increasing the tracking speed of the system. Paralleling optics placed in front of the system should reduce the size of the mirrors required, decrease total system size, and increase the instantaneous field of view.

To increase tracking speed and image processing, another improvement would be to re-implement the tracking algorithms using a smart camera. This smart camera incorporates a CMOS image sensor and a programmable Digital Signal Processor (DSP) to process the image data. The camera and DSP combination will replace the PC for image processing, because framegrabbers are not necessary. The CMOS image sensor allows windowing (pixel addressing) to speed up the frame rate. Using a smart camera, the system will be a stand-alone system for fast motion tracking, because the camera will send control signals directly to the control system.

Because this system can obtain a stereo image through one camera, it solves synchronization challenges and can be used in many applications in addition to mobile robotics where low power and fast tracking is desired.

The low power requirements make this system promising for mobile robotics applications. Our system achieves velocities and accelerations comparable to a current system that has fifty times the power requirements, and it surpasses another popular system that has six times the power requirements.

[1] Metrica TracLabs Biclops PTV head (taken from [5]).
[2] Helpmate Zebra Vergence PTV head (taken from [5]).

References

[1] A. Goshtasby and W. Gruver. Design of a single-lens stereo camera system. *Pattern Recognition*, 26(6):923–937, 1992.

[2] D. Lee and I. Kweon. A novel stereo camera system by a biprism. *IEEE Journal of Robotics and Automation*, 16(5):528–541, October 2000.

[3] F. Martin. *Robotic Explorations: An Introduction to Engineering Through Design.* Prentice Hall, 2000.

[4] S. Nene and S. Nayar. Stereo with mirrors. In *Proceedings International Conference Computer Vision*, pages 1087–1094, 1998.

[5] K. Nickels and D. DiPaolo. Pan-tilt-verge stereo heads for mobile robots: The use of software components in driver design. In *Proceedings of the 5th World Multiconference on Systemics, Cybernetics and Informatics (SCI2001)*, volume VII, pages 455–460, July 22-25 2001.

[6] W. Teoh and X.D. Zhang. An inexpensive stereoscopic vision system for robots. In *Proc. IEEE International Conference Robotics and Automation*, pages 186–189, 1984.

[7] TRACLabs. Introducing biclops. http://www.traclabs.com/tracbiclops.htm Accessed 7/23/2002.

[8] P. Zanen. Single lens 3D with convergence compensation. U.S. Patent Nos. 5,532,777 and 5,883,662.