A Novel Mechanism for Stereo Active Vision

Harley Truong†, Samer Abdallah†‡, Sebastien Rougeaux‡ and Alexander Zelinsky‡

†Department of Engineering
Faculty of Engineering and Information Technology
The Australian National University
Canberra ACT 0200, Australia
http://engn.anu.edu.au

‡Robotic Systems Laboratory
Research School of Information Sciences and Engineering
The Australian National University
Canberra ACT 0200, Australia
http://syseng.anu.edu.au/rsl/

Abstract

This paper deals with the mechanical aspects of robot vision and focuses on the design and implementation of a high-performance system. The system is called CeDAR (Cable Drive Active vision Robot). As part of the research at the Robotic Systems Laboratory (RSL) to develop real-world vision systems, CeDAR is designed for speed and accuracy through its novel use of cables in zero-backlash transmissions and a parallel mechanical architecture. By optimising the design for agility, the system is able to a carry a large payload (700g) while achieving its performance specification of five 90° saccades per second with an angular repeatability of 0.01° in all three axes.

1 Introduction

A brief overview of previously built active vision devices reveals a trend towards smaller, more agile systems. In the past the goals were to experiment with different configurations using large systems with many DOFs, like the KTH active head [Pahlavan and Eklundh, 1994] with its 13 DOFs and Yorick 11C [Sharkey et al., 1993] with a 55cm baseline and reconfigurable joints. More recently, smaller active heads such as the palm-sized Yorick 55C [Sharkey et al. 1998] and Escher [Kuniyoshi et al., 1995] with an 18cm baseline have been designed to mount on mobile robots for active navigation and for telepresence applications.

The trend towards smaller active vision systems comparable in size to the human head is pushing the limit of motor, gearbox and camera design. In most systems, the size of the motors and cameras limit the compactness of the active head and the motors themselves add to the inertia of moving components. A notable exception to this is the Agile Eye [Gosselin et al., 1996] where no motor carries the mass of any other motor. Such a parallel mechanical architecture was the inspiration for the drive system in our active head (Figure 1).

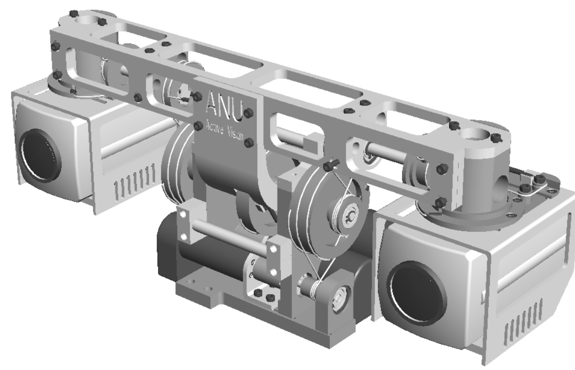

Figure 1: CAD model of a cable-driven active head.

Another issue in the pursuit of faster and more accurate active heads is the choice of transmission system. The need for backlash-free speed reduction is critical for high-speed applications and the most common way this is solved is with harmonic-drive gearboxes. All three versions of Yorick as well as Escher use harmonic-drive technologies. A disadvantage of the technology is an unavoidably large speed-reduction ratio that limits the output speed to less than 100rpm [HD Systems]. This limitation is seldom a problem in applications like smooth pursuit where joint velocities rarely saturate. But during high speed movements like saccades, where the motors are driven at maximum acceleration to travel from one extreme position to the other, velocity saturation is of concern. Cable drive technology is an alternative to harmonic drive gearboxes that does not have speed limitations. The advantages of cable drive are discussed in later sections.

An earlier prototype (Figure 2) built at the ANU Robotic Systems Laboratory proved the usefulness of cable-drive transmissions and parallel mechanical architectures in a 2 degree-of-freedom active 'eye' system [Truong, 1999]. The prototype was fast, responsive and accurate. CeDAR applied the knowledge learnt from the earlier design but in a stereo configuration.

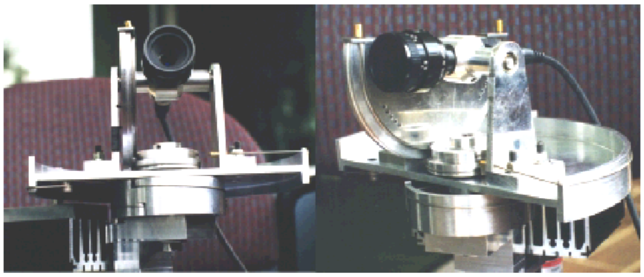

Figure 2: Earlier prototype of a 2DOF active head.

The goals were still to develop a high-performance active head for real-time, real-world applications, but the addition of a second camera brought with it new benefits as well as new challenges. With the complexity of a second camera and a much heavier and larger payload, the margin for error was that much smaller.

This paper documents the design of the CeDAR system from initial performance specifications through to the choice of kinematics, transmission system and mechanical architecture. An overview is given of the hardware components used and the results of the performance testing is presented. Finally, a brief synopsis of future developments is given.

2 Performance Specifications

Table 1 lists the performance specifications for the active head. The maximum range, payload and baseline specifications were based on the desire to use larger motorised-zoom cameras. The saccade rate and pointing accuracy were chosen based on the desired performance of the device in its intended application. Real-time tracking is the desired task and there is a direct relationship between our task-oriented specifications and the minimum requirements for effective tracking [Brooks et al., 1998].

Specification	Tilt axis	Verge axes
Maximum range	90°	90°
Maximum velocity	$600° s^{-1}$	$600° s^{-1}$
Maximum acceleration	$10000° s^{-2}$	$10000° s^{-2}$
Angular resolution	0.01°	0.01°
Angular repeatability	0.01°	0.01°
Saccade rate	5Hz	5Hz
Payload	700g	
Baseline	30cm	

Table 1: Performance specifications

3 Mechanical Design

3.1 Kinematics

There are two widely used configurations for stereo active platforms, the Helmholtz (Figure 3) and the Fick configuration (Figure 4). Each design differs in the order the axes are driven: either tilt first and then verge or verge first and then tilt. Also, the Helmholtz configuration has a common tilt axis. A description of the merits of each design is given in [Murray et al., 1992]. Whatever the configuration, an essential feature is that there must be a common tilt plane between the two cameras in order to implement vergence and stereo matching. The most common is the Helmholtz configuration because mechanical coupling is the easiest way to guarantee a common tilt plane. The more difficult alternative is coupling at the control level. CeDAR is arranged in the more popular Helmholtz configuration with three axes: left vergence, right vergence and a common tilt (elevation) axis.

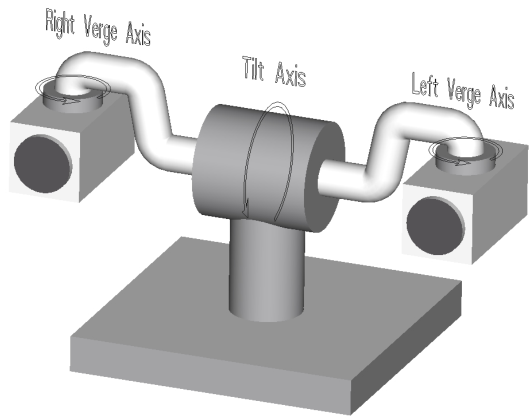

Figure 3: Kinematics of a Helmholtz configured head.

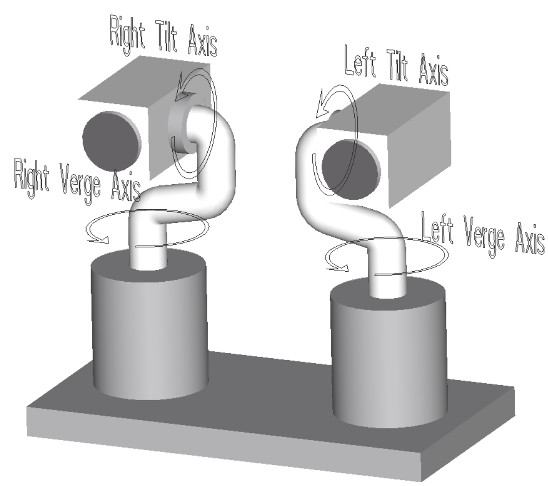

Figure 4: Kinematics of a Fick configured head.

An important kinematic property of the design is that the axes intersect at the optical center of each camera. For vision processing this reduces translational effects and the number of unknown parameters that need calibration. Ignoring these offsets is justified if they are within manufacturing precision and small compared to the distance to the target object [Sharky et al., 1997].

3.2 Transmission System

The transmission system used in CeDAR is the same as the one used in the first prototype which was inspired by a cable driven manipulator [Townsend, 1988]. A cable drive transmission consists of a pulley, a smaller diameter pinion and a cable that wraps around both the pulley and the pinion (Figure 5).

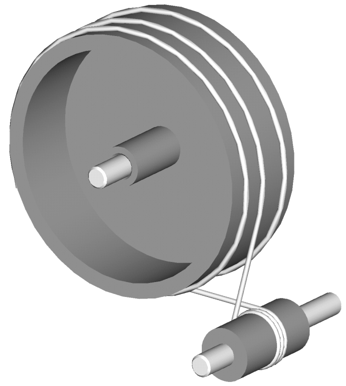

Figure 5: Cable transmission system.

The principle is the same as in gear transmissions except force is transmitted by tension in the cables and not by contact between gear teeth. Speed reduction, similar to gear transmissions, is proportional to the ratio of pulley and pinion diameters. There are many advantages in using cables.

- **No backlash:** force is transmitted by tension in the cables rather than contact forces between gear teeth as in the case of conventional gearboxes. With gearboxes, a small amount of backlash is always necessary to reduce excessive wear, heat and noise. No such requirements exist for cable transmissions.
- **No slippage:** unlike belt drive, the cables are terminated at each end and torque is transmitted to the pinion by several turns of cable to prevent slippage.
- **No lubrication:** the cables do not experience wear or friction like gearboxes and therefore do not require lubrication.
- **High efficiency:** typically 96% [Townsend, 1988] compared to 80% for planetary gearboxes [Maxon Motors].
- **No speed limits**: unlike harmonic gearboxes that are limited to less than 100rpm [HD Systems], cable drive does not have such speed limitations.
- **Torque limited only by strength of cables:** A wide range of cable sizes are available [Sava Cables]. We use a 1.12mm diameter cable with 343 strands and a breaking strength of 77kg.

There are some disadvantages in using cables as compared to conventional gear trains. The first is a finite angular range due to the cables not forming a continuous loop. A typical 20:1 cable transmission has about one revolution of travel at the output stage with the limiting factor being the space needed to wind the cable onto the pinion. Another disadvantage is the difficulty in miniaturizing the transmission. The largest pulley in our design is 6cm in diameter compared to 2cm for typical planetary and harmonic gearboxes. The limiting factor is the minimum bend radius of the stainless steel cables that prevents the use of smaller diameter pinions and pulleys. Future prototypes may use other types of cables like synthetic fibres that have better strength to thickness ratios and more flexibility.

However, in well designed active heads, the disadvantages just mentioned are not relevant because (i) the angular ranges of the joints are limited to 90˚ (Table 1), so the finite travel limitations of cable drive are immaterial and (ii) if the pulleys are integrated into structural members, then the size of the transmission is no longer an issue. For example, in our active head, the final stage bevel is part of the camera mounting bracket (Figure 7).

Figure 6 shows the cable circuits in the active head design. There is a two-stage 25:1 reduction in the tilt joint and two-stage 19:1 reductions followed by bevel transmission in the verge joints. Multistage reductions produce a more compact transmission compared to the bulkier single stage reduction in the earlier prototype. Also, the transmission ratios are optimised and match the source and load inertias to provide maximum acceleration for a given torque input from the motors [Pasch et al., 1984]

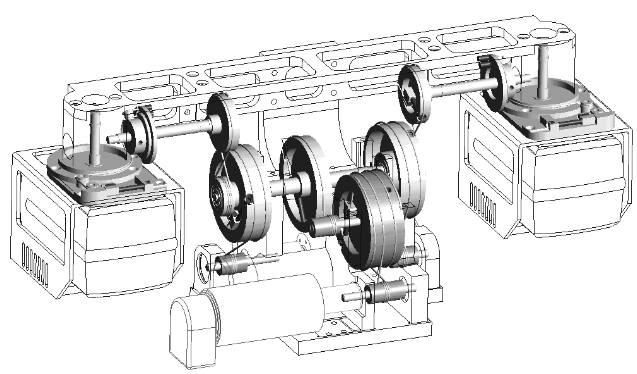

Figure 6: Rear view showing cable circuits.

An interesting part of the cable system is the bevel transmission that transmits torque across orthogonal shafts. The key part of the design is the use of two cables: one for forward motion and one for backward motion. Each bevel has two cable-wrapping surfaces with different diameters so that there are two points of intersection between the bevels for the cables to jump across. If there were only one wrapping surface per bevel, then both cables would have to cross over at the exact same point, which is physically impossible. Figure 7 shows a CAD model of the bevel system revealing the details of the cabling. There are two cables and three terminators (black lugs), one of which is used for tensioning.

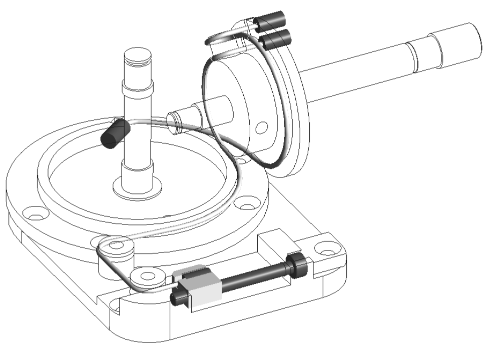

Figure 7: Cable drive equivalent of a bevel gear.

3.3 Mechanical Architecture

Inspired by devices such as the Agile Eye [Gosselin et al., 1996], the active head has a parallel mechanical architecture. Figure 8 shows how all the motors are fixed to the base so that they do not contribute mass to any of the joints. The advantage in doing so as opposed to locating the motors on the tilt joint itself is that the load placed on the tilt motor is lessened. Another advantage is that cable management is easier: the motor and encoder wires do not have to pass through awkward joints to reach the base.

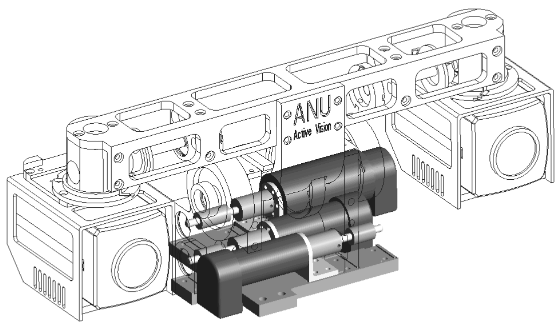

Figure 8: Front view showing all the motors fixed to an immovable base.

The penalty of having a parallel architecture is that it makes the device more complex. Indeed, adding a fourth degree of freedom, a global pan (neck) joint, and still keeping to the parallel drive architecture would be challenging.

3.4 Design Optimisation

To aid the design, simulation and optimisation process, the active head was designed completely within Pro/Engineer [PTC], a parametric solid modeling package. The package allowed a smooth flow of ideas from the initial sketch to the 3D parametric model to the final 2D technical drawings and also the parametric feature allowed design changes to be made by simply typing in the new dimensions. And because the working model was a solid model, material properties could be assigned and mass properties such as inertia, mass, volume and center of gravity were easily obtained. The inertial information was particular useful in optimising the transmission ratio.

The important simulation tools used were the motion and structural simulators. The motion simulator found such things as velocities, accelerations and reaction forces at any point in the model given a trajectory and description of the joints. The structural simulator used finite element techniques to find stress and strain distributions across a component for a given loading condition. Both tools were used extensively to optimise key components. For example, the main structural support holding the cameras was optimised for strength and stiffness. The results of the finite element analysis are shown in Figure 9.

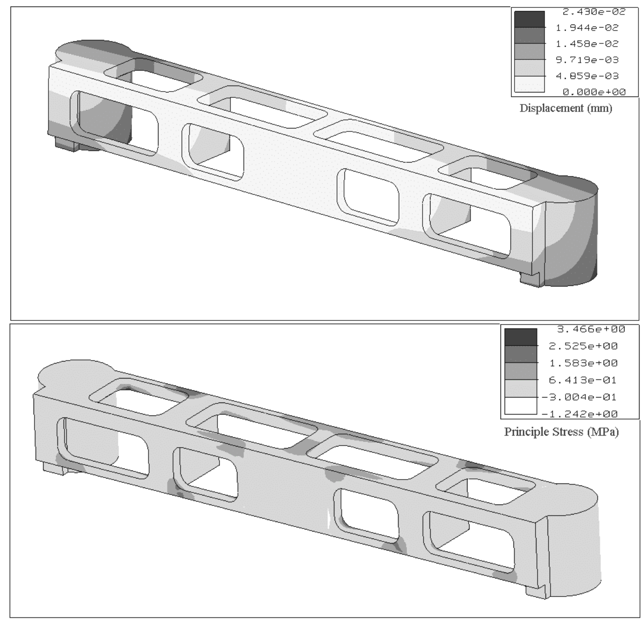

Figure 9: Finite element analysis showing displacement and stress distributions.

4 Hardware Overview

Figure 10 shows the fully assembled active head. Machining time was 200 hours to produce the head, with 80% of the parts produced by hand on a lathe and a mill and 20% produced with a numerically controlled 3-axis mill. The fully assembled head weighs 3.5kg with a moving mass of 1.7kg including the 700g payload.

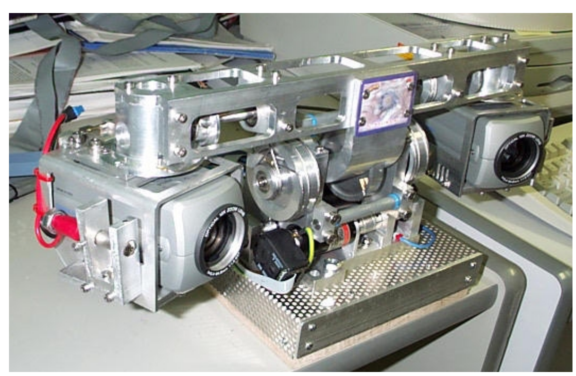

Figure 10: Front view of CeDAR.

A 300Mhz Pentium II computer running Linux is the brain of the system, accepting video images from the digital cameras through a Fire-Wire interface (capable of transmitting at 400 Mb/s), and sending motion commands to the Motion Engineering Inc (MEI) control card through a PCI interface. The control card implements a PID control law using encoder feedback from the motors (1000 pulse/rev resolution). Pulse Width Modulation (PWM) amplifiers are used to amplify the analog signals from the control card before driving each of the three motors in the active head. Each motor has its own 250 watt amplifier and all are powered by a single 600 watt switch-mode supply. Figure 11 illustrates the hardware components in the system.

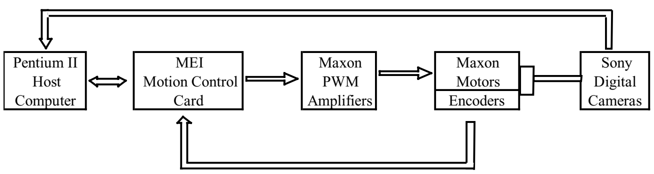

Figure 11: Block diagram of the hardware components.

5 Performance Testing

A software routine was written to test the speed performance of CeDAR by driving the joints to their maximum range, speed and acceleration in a cyclic fashion. In other words, the joints were driven to execute saccades repeatedly. During this test, the command positions, actual positions of the joints and time were logged at millisecond intervals. The position data was then differentiated using a three-point rule and filtered using a 7-point moving average to obtain the velocity and acceleration profiles shown in Figure 12. The graphs show that the vergence joints achieved a maximum saccade rate of 6 per second over a range of 90° and a peak acceleration of 20,000°/s^2 with a maximum velocity of 800°/s. The tilt joint achieved similar performance with a maximum saccade rate of 5 per second over a range of 90° and a peak acceleration of 18,000°/s^2 with a maximum velocity of 600°/s.

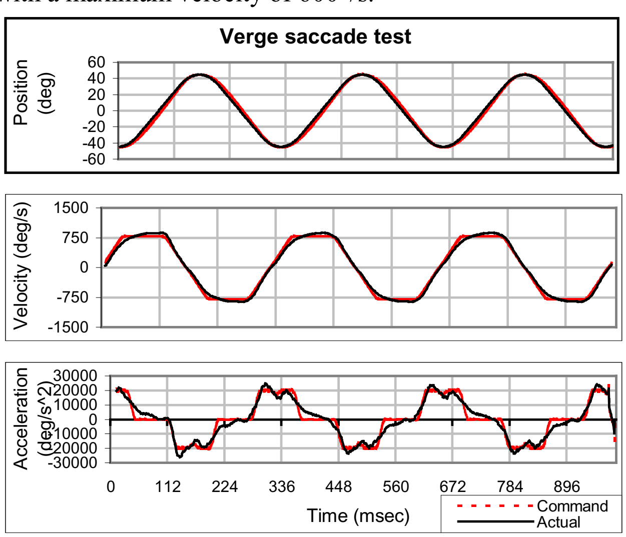

Figure 12: Speed testing plots.

A series of accuracy tests were also conducted using laser pointers mounted to the sides of the cameras. Laser pointers were used instead of the cameras themselves because of the ease and accuracy with which measurements could be taken. A pen and ruler were used to measure the distances between laser patterns on a wall 5 metres away from the active head and then trigonometry was used to convert these linear measurements into angular measurements. The experimental setup is shown in Figure 13.

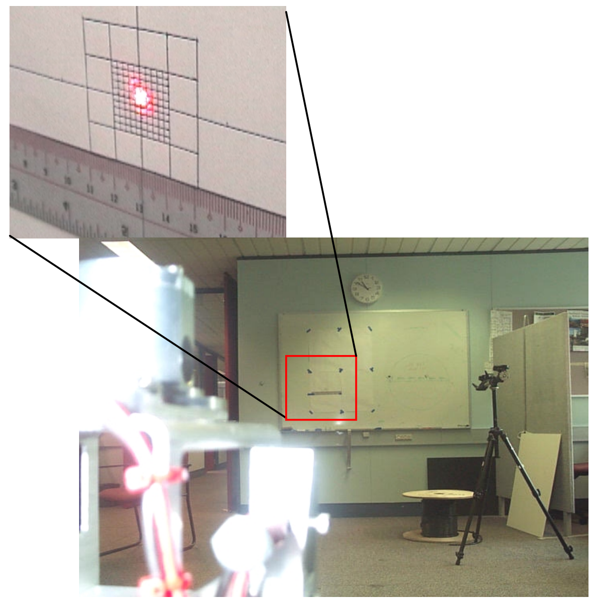

Figure 13: Experimental setup for accuracy measurements.

Three tests were performed to measure CeDAR's accuracy in three respects: angular repeatability, resolution and the ability to perform coordinated motion.

Repeatability: the ability to return to an absolute position after a series of complex movements was demonstrated by moving the joints to an arbitrary position, relocating to another location and then to return to the original point. In systems that suffer from backlash, friction or compliance the return point differs from the original.

Angular resolution: the smallest angle that could be sensed and actuated was measured by moving the joints forwards and backwards a small increment and seeing if the laser point moved accordingly.

Coordinated motion: the test of whether the joints could move in unison both in space and time was demonstrated by verging both laser pointers to the same location on a wall and then commanding the system to follow a predetermined trajectory. Coordination was measured by how closely the lasers were converged throughout the motion.

Table 2 lists the results of the accuracy tests along with the results of the speed tests and the design specifications. All of the specifications were met convincingly thanks to the backlash free operation of the cable drive transmissions.

Specification	Unit	Design Value	Measured Value	
			Tilt	Verge
Saccade rate	Hz	5	5	6
Angular Resolution	deg	0.01	0.01	0.01
Angular Repeatability	deg	0.01	0.01	0.01
Maximum Range	deg	90	90	90
Maximum Velocity	deg/s	600	600	800
Maximum Acceleration	deg/s/s	10000	18000	20000

Table 2: Design and measured performance specifications.

6 Future Work

Applications

As mentioned in section 2, the active head was designed for real-time tracking. Tracking algorithms are in development and will eventually be implemented on the active head. Such applications are based on low-level behaviors like searching, saccading and smooth pursuit as well as the high-level algorithms that 'stitch' these behaviors together. Also, with the benefits of having two motorised-zoom cameras, stereo algorithms could also be used to track objects in 3D-space. An example is for CeDAR to zoom-in on a human face and to follow the person regardless of their movement within the room or to mount CeDAR on a mobile robot for 3D visual navigation.

Hardware Improvements

Most applications in active vision, like tracking and especially mobile navigation require devices with a global pan joint (neck). Further improvements on the active head would implement this feature using a harmonic drive motor. Since the neck joint does not need to move rapidly, there is no need to implement the joint in parallel with the other joints. A simple serial design where the fourth motor would sit beneath the existing head is a straightforward way to do this.

Future Prototypes

An interesting idea that could increase performance and reduce cost and complexity of future prototypes would be to cascade harmonic drive gearboxes with cable drive transmissions. The result would be a high-speed, zero-backlash, compact transmission. Another idea would be to use the Fick configuration to build a head with two independent 'eyes' similar to the pan-tilt device. The advantage in doing so would be to reduce the inertia to essentially only the cameras.

7 Conclusion

There are many approaches in implementing active vision platforms. This paper outlined the novel approaches taken to design a fast and accurate 3 DoF stereo active head. The performance was achieved using cable transmissions and a parallel architecture. Such performance is necessary for real-time applications such as surveillance and navigation and for general purpose systems that are robust and capable of handling real-world environment. This is our ultimate goal and the active head presented in this paper is another step closer to this goal.

Acknowledgments

The authors would like to extend their thanks and appreciation to Dr Jon Kieffer for his initial design of the active head.

References

[Truong et al., 1999] Son Truong, Jon Kieffer, Alexander Zelinsky. A Cable-driven Pan-tilt Mechanism for Active Vision. In Proc. Australian Conference on Robotics and Automation, Brisbane 1999, pg 172-177.

[Pahlavan and Eklundh, 1992] K. Pahlavan and J.O. Eklundh. A head-eye system – analysis and design. CVGIP: Image Understanding: Special issue on purposive, qualitative and active vision, July 1992

[Sharkey et al., 1993] P.M. Sharkey, D.W. Murray, S. Vandevelde, I.D. Reid and P.F. McLauchlan. A modular head/eye platform for real-time reactive vision. In Mechatronics Journal, Vol. 3, No. 4, pp. 517-535, 1993.

[Kuniyoshi et al., 1995] Y. Kuniyoshi, N. Kita, S. Rougeaux, T. Suehiro. Active stereo vision system with foveated wide angel lenses. Asian Conf. on Computer Vision, Singapore, 1995.

[Gosselin et al., 1996] C. Gosselin, E. St-Pierre, M. Gagne. On the development of the agile eye. In IEEE Robotics and Automation Magazine, pages 29-37, Dec, 1996.

[HD Systems] Harmonic Drive Systems. www.hdsystemsinc.com

[Brooks et al, 1998] A. Brooks, G. Dickens, A. Zelinsky, J. Kieffer, S. Abdallah. A High-Performance Camera Platform for Real-time Active Vision. In Field and Service Robotics, pages 527-532, Springer-Verlag, 1998.

[Murray et al., 1992] D.W. Murray, F. Du, P.F. McLauchlan, I.D. Reid, P.M. Sharkey, and M. Brady. Design of stereo heads. In A. Blake and A. L. Yuille, editors, Active Vision, pages 155-172. MIT Press, 1993.

[Sharky et al., 1997] P.M. Sharkey, D.W. Murray, and J. J. Heuring. On the kinematics of robot heads. In IEEE Transactions on Robotics and Automation, vol 13, no. 3, June 1997, pages 437-442.

[Townsend, 1988] W. T. Townsend. The effect of transmission design on force-controlled manipulator performance. PhD thesis, MIT Artificial Intelligence Laboratory, April, 1988.

[Sava Cables] Sava Industries online catalog: www.savacable.com

[Maxon Motor] High Precision Drives and Systems, 1999. www.maxonmotor.com

[Pasch et al., 1984] K A. Pasch, W. P. Seering. On the drive systems for high-performance machines. In Journal of mechanisms, transmissions, and automation in design, Mar 1984, vol 106, pages 102-108.

[PTC] Parametric Technology Corporation. www.ptc.com