# Advancing Active Vision Systems by Improved Design and Control

Orson Sutherland, Harley Truong, Sebastien Rougeaux, & Alexander Zelinsky

Robotic Systems Laboratory
Department of Systems Engineering, RSISE
Australian National University
Canberra, ACT 0200 Australia
http://syseng.anu.edu.au/rsl/

**Abstract:** This paper presents the mechanical hardware and control software of a novel high-performance active vision system. It is the latest in an ongoing research effort to develop real-world vision systems based on cable-drive transmissions. The head presented in this paper is the laboratory's first fully cable-driven binocular rig, and builds on several successful aspects of previous monocular prototypes. Namely, an increased payload capacity, a more compact transmission, and a design optimised for rigidity. In addition, we have developed a simple and compact controller for real-time tracking applications. It consists of two behavioural subgroups, saccade and smooth pursuit. By using a single trapezoidal profile motion (TPM) algorithm, we show that saccade time and motion smoothness can be optimised.

## 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 [6] with its 13 DOFs and Yorick 11C [8] with a 55cm baseline and reconfigurable joints. More recently, smaller active heads such as the palm-sized Yorick 55C [7] and ESCHeR [3] 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 [2] 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).

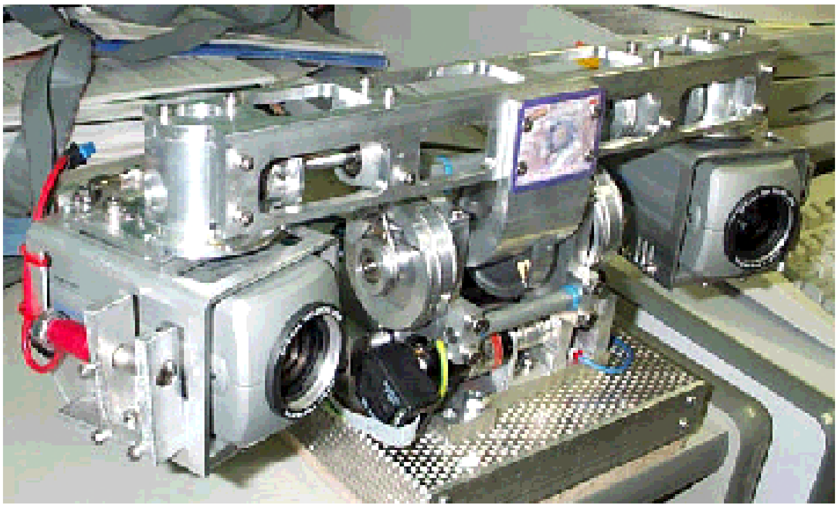
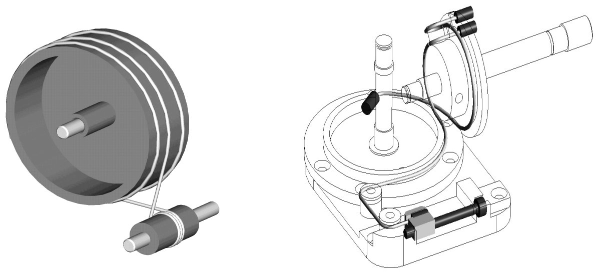


Figure 1. Fully assembled active head.

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 [9]. 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.

Our earlier prototype 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 [11]. The prototype was fast, responsive and accurate. CeDAR applied the knowledge learnt from the earlier design, but in a stereo configuration.

This paper documents the design of the CeDAR system from initial performance specifications through to the choice of kinematics, transmission system and mechanical architecture as well as the hardware components used and the results of performance testing. In addition a control system that makes use of TPM to optimise saccade time and smoothness of pursuit in tracking applications is presented. Finally, a brief synopsis of future developments is given.



a) b)

Figure 2. a) Cable transmission system b) Cable drive equivalent of bevel gear.

## 2. Mechanical Design

### 2.1. Kinematics

There are two widely used configurations for stereo active platforms, the Helmholtz and the Fick configuration. A description of the merits of each design is given in [5]. CeDAR is arranged in the more popular Helmholtz configuration with three axes: left vergence, right vergence and a common tilt (elevation) axis.

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.

### 2.2. Transmission System

The transmission system used in CeDAR is the same as the one used in our first prototype [11]. It was inspired by a cable driven manipulator [10]. 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 2.a).
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.

- **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% [10] compared to 80% for planetary gearboxes [4].

- **No speed limits**: harmonic gearboxes are limited to less than 100rpm [HD Systems], cable drive has no speed limitations.

- **Torque limited only by strength of 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. Another is the difficulty in miniaturizing the transmission. 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), 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 3).

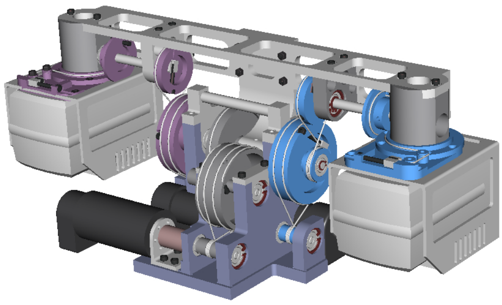


Figure 3. Rear view showing cable circuits.

An interesting part of the cable system is the bevel transmission that transmits torque across orthogonal shafts (Figure 2.b). 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.

### 2.3. Mechanical Architecture

Inspired by devices such as the Agile Eye [2], the active head has a parallel mechanical architecture. Figure 3 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.

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.

Finally, the rig has been optimised for maximum stiffness and minimum weight. This was necessary not only to increase the speed of the head but also its accuracy.

## 3. Hardware Overview

Figure 1 shows the fully assembled active head. It weighs 3.5kg with a moving mass of 1.7kg including the 700g payload.

The video and control hardware consists of Sony digital cameras (DFW-VL500), a Motion Engineering Inc. motion card (PCX/DSP), Maxon DC motors (RE25 and RE36) and a Pentium III computer.

## 4. Performance Specifications and Testing

Table 1 lists the performance specifications for the active head. The maximum range, payload and baseline specifications were based on the potential use of 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 [1].

A software routine was written to test the speed and accuracy of CeDAR. The results are summarised in Table 1. To test speed, the joints were driven to perform repeated saccades. To test accuracy laser pointers were mounted on the sides of the cameras (Figure 1). By programming the head to follow geometric patterns on a wall 5 meters away, we were able to prove **repeatability**, **angular resolution** and **coordinated motion**. Table 1 shows that all specifications were met convincingly. CeDAR's performance compares extremely

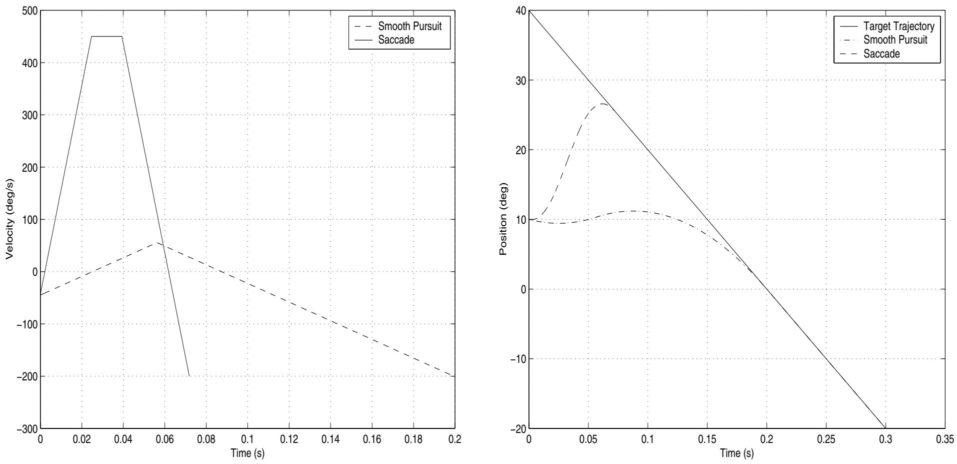
| Specification | Test Tilt | Test Vergence | Spec Tilt | Spec Vergence |
|---|---|---|---|---|
| Max Velocity | $600^{\circ}.s^{-}1$ | $800^{\circ}.s^{-}1$ | $600^{\circ}.s^{-}1$ | $600^{\circ}.s^{-}1$ |
| Max Acceleration | $18000^{\circ}.s^{-}2$ | $20000^{\circ}.s^{-}2$ | $10000^{\circ}.s^{-}2$ | $10000^{\circ}.s^{-}2$ |
| Saccade Rate | $5Hz$ | $6Hz$ | $5Hz$ | $5Hz$ |
| Ang Repeatability | 0.01° | 0.01° | 0.01° | 0.01° |
| Ang Resolution | 0.01° | 0.01° | 0.01° | 0.01° |
| Max Range | 90° | 90° | 90° | 90° |

Table 1. Performance specifications and test results

well to existing heads in addition to its ability to carry a wide range of payloads. Table 2 compares CeDAR's peak vergence velocity and acceleration to two key designs.

| Specification | CeDAR | ESCHeR | Agile Eye |
|---|---|---|---|
| Max Velocity | $800°.s^-1$ | $400°.s^-1$ | $1000°.s^-1$ |
| Max Acceleration | $20000°.s^-2$ | $16000°.s^-2$ | $20000°.s^-2$ |

Table 2. Comparison with two leading designs.




a) b)

Figure 4. a) Velocity profiles b) Position profiles, during trapezoidal motion for saccade and smooth pursuit

## 5. Control

CeDAR's controller is an extension of preliminary work undertaken by [5] on TPM. In particular our approach allows for the implementation of a single algorithm for both saccade and smooth pursuit, enhancing the simplicity and compactness of the controller's design.

### 5.1. Trapezoidal Profile Motion

The essence of the TPM problem is to catch a target initially a distance $x_0$ from the image center either in the shortest time possible, in the case of saccade, or in the smoothest way possible, in the case of smooth pursuit. Both the joints' and the target's starting velocities are potentially non-zero and disparate. Specifically it causes an axis to accelerate at a constant acceleration to a pre-calculated ceiling velocity[1], coast at this velocity for a given period and then decelerate at the same constant rate as the initial acceleration until the target velocity is reached (see figure 4.a). From a mathematical perspective it is a 4 dimensional problem, where the bang acceleration $a$, ceiling velocity $v$, move time $T$ and total distance travelled $x$ are the unknowns. The initial joint velocity $v_1$, the target velocity $v_2$ and the target's initial distance from the image center $x_0$ are the givens.

[1]The maximum absolute velocity of the TPM trajectory

**The Algorithm**

Since the acceleration $a$ is assumed to be constant, the time taken by the head to accelerate from its initial velocity $v_1$ to the ceiling velocity $v$ is

$$T_a = \frac{s \cdot v - v_1}{s \cdot a}, \tag{1}$$

where $s$ is positive if the head accelerates from $v_1$ to $v$ and negative if it decelerates.

Similarly the time taken to decelerate from the ceiling velocity to the target velocity $v_2$ is

$$T_d = \frac{s \cdot v - v_2}{s \cdot a}. \tag{2}$$

Note that the rate of deceleration is equal to the rate of acceleration.

If $T_c$ is the time spent coasting at the ceiling velocity, the total time of the trapezoidal profile motion is

$$T = T_a + T_c + T_d. \tag{3}$$

The distance traveled by the head during the action is

$$x = \frac{s \cdot v + v_1}{2} T_a + s \cdot v T_c + \frac{s \cdot v + v_2}{2} T_d, \tag{4}$$

but can also be considered as the sum of the initial distance of the target from the foveal center $x_0$ and the distance travelled by the target during the move

$$x = x_o + T v_2. \tag{5}$$

With these general TPM equations the specifics of saccade and smooth pursuit can now be developed.

**Saccade**

Saccades involve changing the head's current position and velocity state to that of the target, as inferred by its previous states, in the shortest time possible (see Figure 4.b). Motion smoothness is not a concern and hence acceleration is set to its maximum possible magnitude. Two cases can arise:

- The ceiling velocity required for the action is less than the maximum allowed velocity and hence no time is spent coasting.
- The theoretical ceiling velocity required for the action is greater than the maximum allowed velocity and hence some time must be spent coasting (see Figure 4.a)

It is useful to assume $T_c$ as initially being zero so that (1), (2), (4) and (5) yield

$$s \cdot v = v_2 \pm \frac{1}{2}\sqrt{4 s x_0 a - 2(v_1^2 + v_2^2) - 4 v_1 v_2}, \tag{6}$$

where the smaller of the two values is taken if $v_2$ is greater than $v_1$ and vice-versa. But if both of these values are in excess of the maximum allowed velocity, acceleration and velocity in (1), (2), (4) and (5) are replaced by their respective maxima from which

$$T_c = \frac{1}{v_2 - s \cdot v}((\frac{s \cdot v + v_1}{2} - v_1)T_a + (\frac{s \cdot v + v_2}{2} - v_2)T_a - x_0), \qquad (7)$$

is calculated and hence $T$ is deduced.

Equation (6) also defines the value of $s$. In particular it must be such that the operand of the radical is greater or equal to zero, hence:

$$\begin{cases} \text{s = +1} & \text{if } (v_1 - v_2)^2 + 4x_0 a \geq 0 \\ \text{s = -1} & \text{otherwise} \end{cases}$$

**Smooth Pursuit**

In smooth pursuit we wish to move from one position and velocity state to the next in a given amount of time with the optimal smoothness (see figure 4.b). To achieve this the acceleration in moving to and from the ceiling velocity must be as small as possible. Again both the coasting and no-coasting cases mentioned above are relevant and again we start by assuming that the coasting time is zero initially, from which (1), (2), (4) and (5) yield:

$$v = \frac{x}{T} \pm \frac{1}{2T}\sqrt{4x^2 - 4Tx(v_1 + v_2) + 2T^2(v_1^2 + v_2^2)}. \qquad (8)$$

If these values are in excess of the maximum allowable velocity of the head, the time constraint is unrealiseable. In this eventuality CeDAR's controller has been implemented to initiate a saccade.

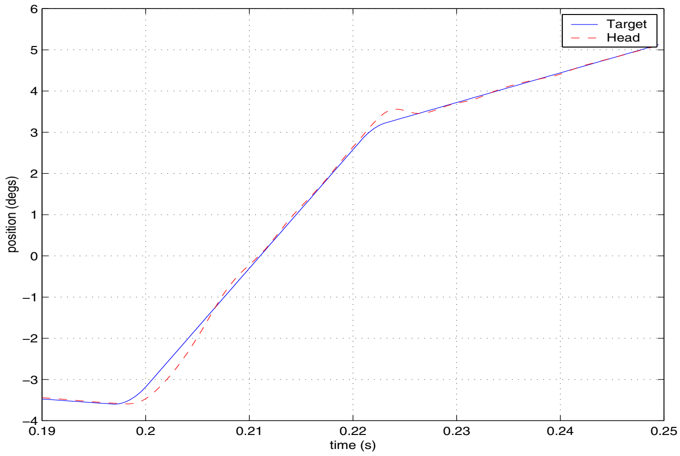
## 6. Future Work

### 6.1. Applications

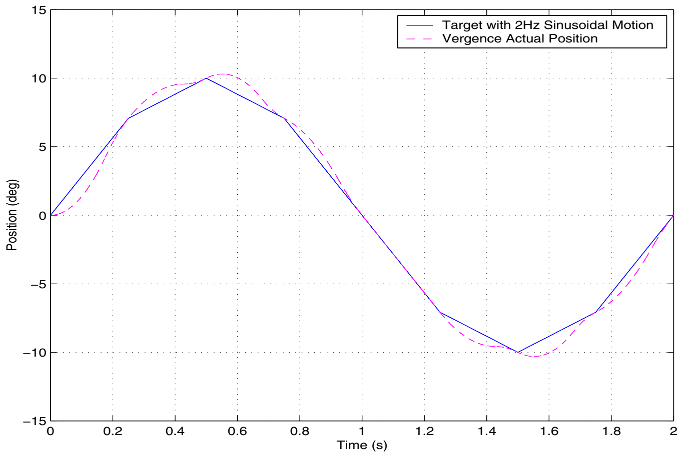
As already mentioned, CeDAR's mechanical and control architectures were designed for real-time tracking. Zero-Disparity filtering and Optic Flow algorithms are in the process of being integrated into the system. Coupled with the TPM controller this should allow for robust tracking. In particular, we intend to demonstrate CeDAR's ability to locate and track a tennis ball during a tennis match.

### 6.2. 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.




a)




b)

Figure 5. a) Two successive real-time saccades b) Smooth pursuit of a 0.5Hz sinusoid, sampled at 4Hz

### 6.3. Future Prototypes

The use of plastics and other polymers in a cable-drive system should see a significant reduction in size and weight. It would also be a low cost way to manufacture moderate to high numbers of active heads. 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

This paper has outlined a novel approach to the design of a fast and accurate 3 DOF stereo active head. Performance was achieved using cable transmissions

and a parallel architecture. Such performance is necessary for real-time applications such as surveillance, navigation and human/robot interaction.
In addition, a simple and compact controller was presented that allowed both saccade and smooth pursuit to be performed for tracking applications. It made use of a single TPM algorithm to optimise saccade time and motion smoothness during smooth pursuit.

## Acknowledgments

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

## References

[1] A. Brooks, G. Dickens, A. Zelinsky, J. Kieffer, and S. Abdallah. A high-performance camera platform for real-time active vision. *Field and Service Robotics*, pages 527–532, 1998.

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

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

[4] Maxon Motors. High precision drives and systems. www.maxonmotor.com, 1999.

[5] D.W. Murray, F. Du, P.F. McLauchlan, I.D. Reid, P.M. Sharkey, and M. Brady. *Active Vision*, pages 155–172. MIT Press, 1992.

[6] 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.

[7] P.M. Sharkey, D.W. Murray, and J. J. Heuring. On the kinematics of robot heads. *IEEE Transactions on Robotics and Automation*, 13(3):437–442, June 1997.

[8] 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. *Mechatronics Journal*, 3(4):517–535, 1993.

[9] HD Systems. Harmonic drive systems. www.hdsystemsinc.com.

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

[11] S. Truong, J. Kieffer, and A. Zelinsky. A cable-driven pan-tilt mechanism for active vision. In *Proceedings Australian Conference on Robotics and Automation*, pages 172–177, Brisbane 1999.