The ISR Multi-Degrees-of-Freedom Active Vision Robot Head

Jorge Batista, Jorge Dias, Helder Araújo, A. T. Almeida

Institute of Systems and Robotics (ISR), and Electrical Engineering Department, University of Coimbra,
3000 Coimbra - PORTUGAL;
phone: +351-39-34876/34884 ; fax: +351-39-35672 ; e-mail: batista@mercurio.uc.pt

Abstract - Experiments in active vision require having the ability to manipulate the visual parameters. The central issue in devellopìng a Multi-Degrees-of-Freedom (MDOF) Active Vision Robot Head is the design strategy of the system. Bringing all the issues of such a system and their solutions together and build a head-eye system with reasonable performance vs. cost is an engineering problem.
This paper presents the main aspects of the ISR MDOF active vision system, design, performance, control and the architecture of the all system, including mechanical and optical degrees of freedom.
To be able to effectively use multi degree of freedom (MDOF) camera systems we need to know how variations in the camera's control parameters are going to cause changes in the produced images. For this we need to have good mathematical models describing the relationships between the control parameters and the parameters of the resulting images, i.e., we need to calibrate the system.
The calibration of the system includes two parts: the *camera calibration problem*, i.e., the calibration of the intrinsic and extrinsic parameters of the camera, and the so-called *kinematic calibration* to calibrate relationships (rotation and translation) between different systems. In this paper only the camera calibration problem is addressed. A method for computing the camera parameters by tracking features in the image when the camera undergoes pure rotation is used.

Keywords - Active Vision, Robot Heads, Systems Architectures, Active Calibration

1. Introduction

Experiments in active vision requires the ability to manipulate the visual parameters. This ability and associated issues form the subject of this paper.

The central issue in developing a Multi-Degrees-of-Freedom (MDOF) Active Vision Robot Head is the design strategy of the system. This problem in our work was formulated as : "how should a head-eye system be designed, what are the design criteria, how and in accordance with what strategy should the head be designed and controlled, what kind of degrees of freedom must be included ?".

The design of a MDOF head-eye system for active vision is dependent on what we put into this notion. First of all, an active vision system is not just an optomechanical device feeding a computer and carrying out the commands from the computer. The degree of integration is crucial for such a system, and of course the issue of real-time processing and control. These factors determine the behaviour one can obtain. The more elaborately the visual system reacts to the surrounding environment, the more evolved the primary tasks will be. The nature of the visual process we want to integrate will be related to the architeture chosen.

One important aspect in the designing stage of these robotic systems is the performance they should accomplish. The analisys of some characteristics of the human active visual system can be usefull for determining performance requirements for velocity and acceleration of a mechanical device that intend to simulate the human visual system behaviour.

Much work has been done in devellopìng vision systems to study how these and other features of the visual system are used to facilitate perception (see [7] for an overview). One of the earliest active vision systems was built by Krotkov et al. [23] at the University of Pennsylvania and the system included two cameras with computer controlled motorized lenses, pan and tilt rotation of both cameras, two translational degrees of freedom and coupled symmetric vergence. Many other researchers have assembled commercial motion-control components such as rotational and translational stages, providing small systems with good repeatability, high-speed motions, but with some restritions in terms of the degrees of freedom available [21;26]. In the late of the eighties, Poggio et al. [22] devellopped at the MIT a vision system where two cameras with motorized lenses are rigidly attached to a mobile platform, and the camera movement is achieved indirectly by pivoting a front surface mirror mounted in front of each camera. At the University of Rochester and Harvard University, binocular active vision systems have also been devellopped. The Rochester active vision system [24] was mounted on a six-degrees-of-freedom robot arm and it has independent vergence axes and coupled tilt movement for both cameras. No motorized lenses has been used. The Harvard system [25] was mounted on a mobile platform and the head itself performes pan, tilt and anti-symmetric vergence motions to control the orientation of the cameras, and focus and aperture control for accomodation of the optical system. The Yorick head developed at the Oxford University [28] as well the Triclops devellopped at the NIST [27] are based on the same mechanical structure, including pan, tilt and independent vergence, being the Yorick equipped with independent camera tilt movement, and the Triclops equipped with a third low resolution camera located at the neck rotation center. Both of these systems presents remarkable performances in terms of speed and accuracy.

The KTH head developed at the Royal Institute of Technology in Sweden [10] was one of the first systems completely motivated by biological reasons, and it includes 13 degrees of freedom, with two mechanical degrees of freedom for each of the two eyes, pan and tilt for the neck, baseline control, and three optical degrees of freedom for each of the eyes. One interesting feature of this system is the optical center adjustment degree of freedom that allow the translation of the eye (camera and lens) in order to compensate the drift of the optical center when focus and zoom movements are performed. This degree of freedom keeps the optical center near the vergence rotation center, despite changes in optical parameters.

The ISR MDOF active vision robot head is probably the head that currently has more degrees of freedom. In addition to the common degrees of freedom for camera heads (pan, tilt and independent vergence for each of the eyes), this head includes the swing movement of the head neck, baseline control, cyclotorsion of the lenses and the ability of adjusting the optical center of the lenses. The mechanical structure of this head is quite similar to the KTH head structure, and just like the KTH head, biological reasons played the main role in the design strategy.

In spite of all the performances and degrees of freedom of the active vision system, to be able to effectively use these systems we need to know how variations in the camera's control parameters are going to cause changes in the produced images. For this we need to have good mathematical models describing the relationships between the control parameters and the parameters of the resulting images, i.e., we need to calibrate the system.

Recently, a lot of emphasis has been placed on algorithms that do not require any camera calibration and on camera calibration techniques that allows the camera to calibrate itself as it moves in an unstructured world.

> *"Eyes in humans and other animals do not need any artificial assistance for calibration".*

The answer to the above observation cames with the fact that eye movements simplify the calibration problem.

Methods for computing the camera parameters by tracking features in the image, without using special patterns for calibration, have been developed by Faugeras [19], Hartley [16], Dron [11], McLauchlan [14], Brady [5], Stein [6] and Basu [2].

We based our approach on the use of feature correspondences from a set of images where the camera has undergone *pure rotation*. The closer the features are located to the camera the more important it is that the camera does not undergo any translation during the rotation. A method to ensure that the axis of rotation passes close to the center of projection (front nodal point in a thick lens model) is presented.

An important aspect for the MDOF optical systems calibration is the fact that the parameters are changing from time to time, which requires a real-time calibration of the parameters or a pre-calibration of these parameters to build up a look-up-table. To be able to adjust the intrinsic parameters in real-time, we modelize the underlying behaviour of the camera for a large group of different setups, changing the intrinsic parameters on these setups. We use bivariate cubic polynomials to model the relationships between the camera's control parameters and the parameters of the resulting images, such has image magnification, focus distance, optical center adjustment, etc. . The calibration involves performing a least square fit of the model to collected data to determine the coefficients of best fit for the bivariate polynomials.

For the extrinsic parameters (pose estimation of the camera), and since the movements of the head-eye system are controlled, we know how much it has been moved relatively to some initial position. With this procedure, the calibration of the extrinsic parameters must only be done at the initial position, and real-time calibration of the extrinsic parameters can be performed.

	Eye Pan/Tilt	Neck Pan	Neck Tilt
Range of Motion	±45°	±80°	$+90° up$ $-60° down$
Peak Acceleration	$35000°/s^2$		
Peak Velocity	$600°/s$		
Interocular distance	$\sim 64.0mm$		
Foveal-Peripheral resolution ratio	10 : 1		

Table 1. Human active visual system characteristics

This paper is divided in two major parts: first we are going to describe all the aspects of the ISR MDOF active vision system, namely the design stategy, performance, control and system architecture, and in the second part we will describe the optical calibration of this system taking advantage of the controled movements that can be done with this system.

2. The ISR-Coimbra MDOF Active Vision Head

Active vision systems are often modeled on attributes of the human visual system since this is the most well-studied visual system. The human oculomotor is one of the best known functions of our brain. Two attributes of the human vision system, ocular motion and foveal-peripheral vision, are essential to human visual perception. Ocular motion allows movements of the eyes to direct the view point of the visual system. Foveal-peripheral vision enables humans to perceive small regions in fine detail in combination with a wide field of view at coorse detail. Taking advantages of the ocular visual system, the human head also has the capability of changing gaze and fixate on features of the environment.

Our main purpose for building these vision robot head was not just having an active vision system with the basic degrees of freedom (vergence, pan and tilt) to perform tracking or to be used on mobile platforms, but also to have a device where we can study and simulate some of the human visual behaviours.

One important aspect in the designing stage of these robotic systems is the performance they should accomplish. The analisys of some characteristics of the human active visual system can be usefull for determining performance requirements for velocity and acceleration of a mechanical device that intend to simulate the human visual system behaviour.

Some characteristics of the human visual system are sumarized in table 1, being the information presented in the table obtained from Carpenter [15], Yarbus [1], Geiger and Yuille [3], Webb Associates [20] and Wurth and Goldberg [18].

2.1 Mechanical Structure

The ISR MDOF active vision robot head has the following mechanical degrees of freedom and some characteristics of the mechanical degrees of freedom are summarized in table 2:

Eyes-mechanical Each eye has three degrees of freedom (a total of six):

- elevation (tilt)
- azimuth (pan)
- cyclotorsion (being developed)

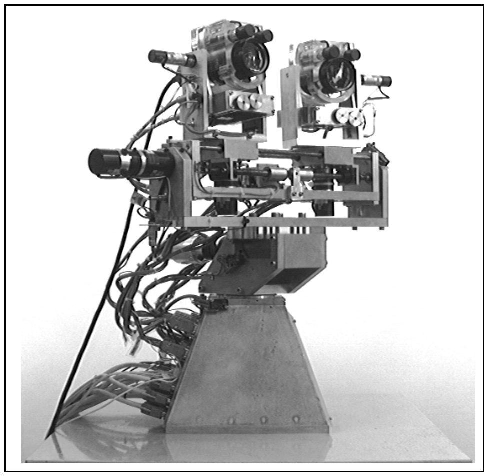

Fig. 1. ISR MDOF Active Vision Head

	Precision	Range	Velocity
Neck Pan	0.0036°	$[-110°..+110°]$	$\sim 360°/s$
Neck Swing	0.0036°	$[-27.5°..+27.5°]$	$\sim 360°/s$
Neck Tilt	0.0036°	$[-32°..+32°]$	$\sim 360°/s$
Eye Pan	0.0036°	$[-45°..+45°]$	$\sim 360°/s$
Eye Tilt	0.0031°	$[-20°..+20°]$	$\sim 330°/s$
Cyclotorsion [1]	0.0031°	$[-25°..+25°]$	$\sim 330°/s$
OCA [2]	$8nm$	$[0..80]mm$	$\sim 1mm/s$
Baseline	$20nm$	$[137..287]mm$	$\sim 5mm/s$

Table 2. Mechanical structure characteristics of the ISR MDOF active vision system

- an aditional degree of freedom is included to keep the optical center at the crosspoint of the azimuth and elevation axes of the lens.

Neck-mechanical The neck has three degrees of freedom:

- tilt
- pan
- swing or lateral tilt movement

Baseline The ability of mechanicaly change the distance between the two eyes.

The ISR MDOF active vision robot head is probably the head that currently has more degrees of freedom. In addition to the common degrees of freedom for camera heads (pan, tilt and independent vergence for each of the eyes), this head includes the swing movement of the head neck, independent tilt movement for both eyes, and the ability of adjusting the optical center of the lenses. The latter is to ensure pure rotation when verging the cameras and compensate for the translation movement of the optical center when changing the focal length of the lens. Cyclotorsion of the eyes is at the moment being developed.

This design of the head inspired by biological motivation has direct consequences on the kinematics of the head. No coincident axes have been possible for all the three neck degrees of freedom. Only the pan and swing axes intersect. The tilt axis does not intersect any of these two axes and it was put 8 cm ahead and 14 cm above of the pan and swing axes. With this particular design the eyes will have a translational component added to the pan and swing rotation movement. Due to this mechanical design a much more demanding kinematics calibration and control of the head are required.

The eyes of this head are equipped with fully independent movements and azimuth, elevation and cyclotorsion are available. The inclusion of independent neck and eyes elevation movements was motivated by the fact that a smooth-pursuit of light loads is accomplished with much more accuracy and saccadic movements of the eye can be performed much faster than neck saccadic movements. We included the optical center adjustment due to the fact that this head is equiped with motorized zoom lenses. Pure rotation vergence movements are possible using this degree of freedom. We don't think that the adjustment of the optical degree of freedom is crucial for active vision robot head, but the kinematic of the eye becomes a lot easier, in special for motorized zoom lenses. Pure rotation is also important to implement distance-independent saccade algorithms, and is essential for algorithms that assume that the relationship between motion space and motion in joint space may be learned without knowledge of the target distance. This could be extremely important for example to perform active calibration of the optical degrees of freedom, as it will be show later. The optical center adjustment only takes place along the optical axis of the lens, since the larger variation of the center of projection occurs along this axis as a result of focus and zoom changes. A small variation on the location of the center of projection also occurs on the other two axes, but we considered that variation negligible compared with the variation that occurs along the optical axes. The eye mechanical structure of this head and its optical center adjustments constraints limits the type of lenses that can be used. The mechanical structure develloped for the cyclotorsion movement of the eye limits also the type of camera that can be used. Only small cameras can be used, not only because of its weight but also due to its size.

The dynamic performance, accuracy, and other requirements are achieved with harmonic drive DC motors. In order to simulate the performances of the human visual system there is a requirement for large acceleration, low friction, high repeatability and minimal transmission errors. These are some of the primary characteristics of systems that use motors and feedback devices mounted directly to the axes of motion. With the harmonic drive gearboxes, transmition compliance and backlash, which can cause inaccuracy and oscilations, are almost eliminated. Unfortunately, in order to include all the degrees of freedom described, it was impossible to reduce the loads on the axes only to inertial loads. This imposes the use of higher torque DC motors. All the motors are equipped with optical encoders that provide good resolution but requires initialization procedures each time the system is powered up.

2.2 Optical Structure

In real world environment the range of conditions that a camera may need to image under, be it focused distance, spacial detail, lighting conditions or radiometric sensitivity, can often exceed the capabilities of a camera with a fixed parameters lens. To adapt the imaging conditions the camera system requires lenses whose intrinsic parameters can be changed in a precise and fast controlable manner.

Motorized lenses offer greater capability and flexibility than fixed-parameter lenses, however, most active vision systems have been limited to cameras with fixed lenses because of the difficulty of modeling cameras with motorized lenses, their weight and the precision they offer. Nowadays, mo-

---

[1] Being developed

[2] OCA : Optical Center Adjustment

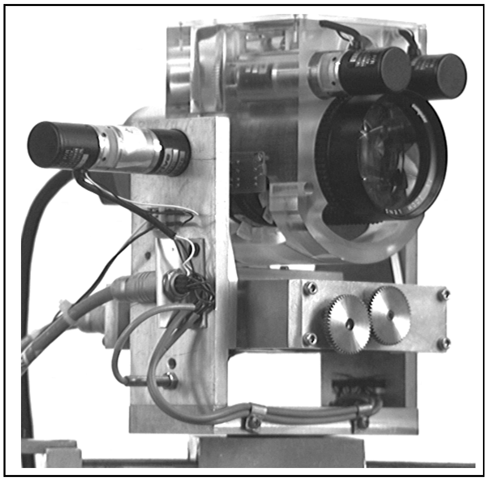

Fig. 2. ISR MDOF Eye with the ISR MDOF Motorized Zoom lens

	Precision	Range	Velocity
Zoom	$Range/90000$	$[12,5..75]mm$	$\sim 1.2 * range/s$
Aperture	$Range/50000$	$[1,2..16]$	$\sim 2.2 * range/s$
Focus	$Range/90000$	$[1..\infty]m$	$\sim 1.2 * range/s$

Table 3. Optical structure characteristics of the ISR MDOF active vision system

torized zoom lenses becames more and more important in active vision systems, e.g., for depth reconstruction, magnification, focusing, etc.. Zoom can be used to acquire images at different magnification, e.g., simulate foveation and concentrate the view on a particularly feature, focus can be used to automatically refocus on objects at different distance and compute relative depth maps, and the aperture can be used to automatically adjust the iris according to the changes in lighting conditions.

Most of the existing heads uses standard motorized lenses with potenciometers as feedback information. These lenses has the disadvantage of moving too slow for real-time acommodation purposes (5-6 seconds to full range movement), and the accuracy for position control is not very good due to the type of information they provide as feedback. New motorized lenses have been developed to enable this head to acommodate the optical system in almost real time, with very good precision (see fig. 2). These lenses have controllable zoom, focus and iris and they use small harmonic drive DC motors with encoder feedback information. For these lenses we used zoom lenses from Computar that we motorized (see fig. 2). By using DC harmonic drives we are able to span the full range of zoom (12.5 mm to 75 mm) and focus (1 m to infinity) in 0.8 sec and the full range of the iris in 0.45 sec.. Also the full range of focus and zoom are discretized into 90000 positions whereas the full range of iris is discretized into 50000 positions (see tab. 3). This motorized zoom lens weight almost 1kg and its dimensions are 8.0cm x 10.5cm x 10.0cm [W x H x D]. With this kind of design we hope not only to fully characterize the operation of this kind of lenses (in terms of the trajectory of the optical center) and calibrate them, but also to try them in applications where this type of speed is crucial.

With such such performances, the lens is able to make continuous, small optical adjustments required by many algorithms in near real time with excellent precision. Qualitative improvements in lens perfomences increase the advantage afforded by active vision techniques that rely on controlled variations of intrinsic parameters.

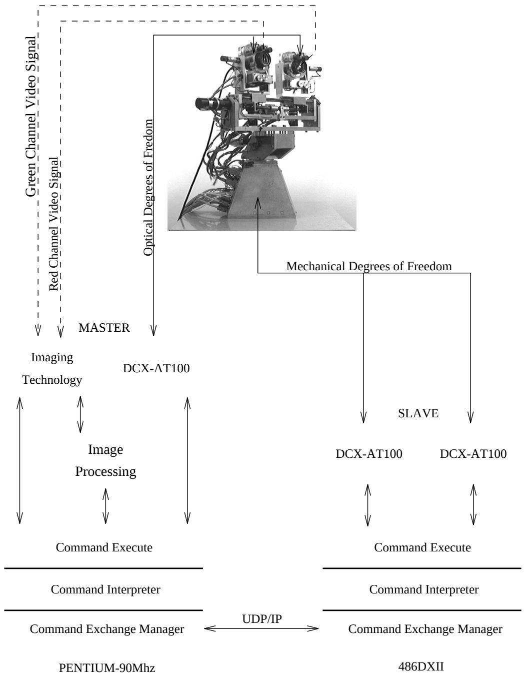

Fig. 3. The ISR MDOF system Architecture

2.3 System Architecture

The ISR MDOF active vision robot head is connected to a pair of PC's being one dedicated to the control of the mechanical degrees of freedom and the other dedicated to the control of the optical degrees of freedom and image aquisition (see fig. 3). This later PC is a Pentium (90Mhz) running as a *Master* being the other PC the *Slave* unit. These control units are connected between eachother through an Ethernet link. A special protocol for commands exchange as been developed using the UDP/IP protocol.

A modular multi-axis motion controller from Precision Micro Control (DCX-AT100) was used to control all degrees of freedom of the head. This modular system consists of a motherboard where up to six daughterboards or modules can be connected, and it is based on a 32-bit 80960 RISC CPU with floating point math processor. On-board *Multitasking* executes up to 10 independent programs or background tasks simultaneously without interrupting motion control. The DCX-AT100 motherboard performs two key functions. First it provides both a physical structure to connect the modules mechanically and electrically. Second, it contains computer logic which is used for controlling and communicating with the individual modules. Another feature is that multiple DCX's can be built into a single system when the application requires more than six modules, just like the ISR MDOF robot head. The computer logic in the DCX system is programmed to interpret and execute commands sent to it by the user. The ability to combine sequences of commands (macro commands) provides the user with a powerful tool for implementing various controls requirements.

Three control boards from Precision Micro Control are used to control the 18 degrees of freedom of the robot head. The DC servo controller module (DCX-MC200) that was pluged in on the motherboard contains a trapezoidal velocity profile generator and a digital PID compensation fil-

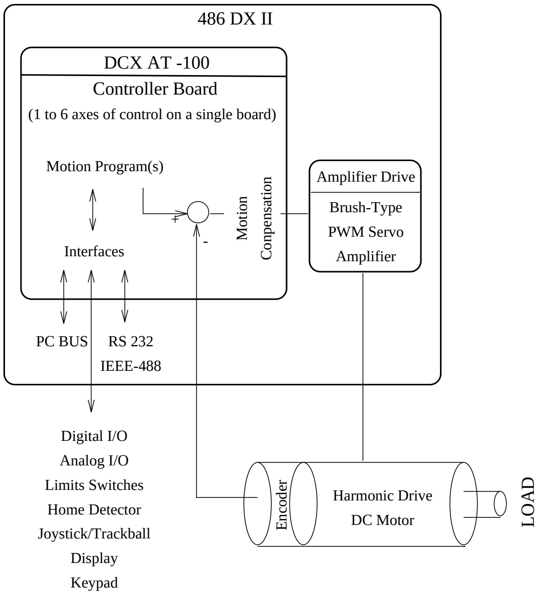

Fig. 4. Motion Control System

ter. Each module is a self-contained intelligent controller. Several control signals are available, including limit switchs, home detector and drive fault (see fig. 4). A low-level motor control interface is provided by the motherboard, being the command exchange between the user and the board realized through a ASCII command management.

The image aquisition is done by an Imaging Technology True Color PC board, being each of the monochrome Cohu 4990 video cameras connected to the red and green channel of the RGB input of the board. All the image processing is at the moment taking place at the Pentium Master PC.

The command exchange protocol developed to connect the Master and Slave units uses three levels of protocol. At the lowest level the *command exchange manager* is responsible for sending and receiving packets of data through the UDP/IP protocol. The *command interpreter* establishes the connection between the lowest and the highest level of the protocol, analising the packets of data and deciding to which machine should the command be sent. This level is only active at the Master unit. At the higest level is the *command execute* that is responsible to send the commands to the board controller.

3. Optical Calibration : Mathematical Background

3.1 The Rotation Method - Theory

The main purpose of this pure rotation calibration procedure is to find camera parameters which will enable one to best predict the effects of camera rotation in some optimal manner. Given any pair of images where the camera has undergone pure rotation over some axis, if the intrinsic camera parameters and the angle and axis of rotation are known one can compute where some feature points from one image will appear in the second image after rotation. This is the main observation that allows us to use pure rotation to obtain some of the intrinsic camera parameters.

Let us assume that the camera is rotated in a rigid world environment around some axis (see fig. 5). Also assume the existence of a camera coordinate system located at the lens optical center and the Z axis viewing along the optical axis of the lens. A 3D point $P_c(x_c, y_c, z_c)$ in the camera coordinate system will move after rotation to a point $P'_c(x'_c, y'_c, z'_c)$ through the matricial relationship

$$P'_c = R \cdot P_c = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} P_c. \qquad (1)$$

Using the perspective projection pin-hole geometry the 3D camera point P'_c projects to the undistorted image point $p'_u(x'_u, y'_u)$ where

$$x'_u = f\frac{x'_c}{z'_c} = f\frac{r_{11}x_c + r_{12}y_c + r_{13}z_c}{r_{31}x_c + r_{32}y_c + r_{33}z_c} \qquad (2)$$

$$y'_u = f\frac{y'_c}{z'_c} = f\frac{r_{21}x_c + r_{22}y_c + r_{23}z_c}{r_{31}x_c + r_{32}y_c + r_{33}z_c}. \qquad (3)$$

Multiplying equations 2 and 3 by f/z_c and substituting $x_u = f(x_c/z_c)$ and $y_u = f(y_c/z_c)$ results

$$x'_u = f\frac{r_{11}x_u + r_{12}y_u + r_{13}f}{r_{31}x_u + r_{32}y_u + r_{33}f} \qquad (4)$$

$$y'_u = f\frac{r_{21}x_u + r_{22}y_u + r_{23}f}{r_{31}x_u + r_{32}y_u + r_{33}f}. \qquad (5)$$

Observing these last two equations, we can conclude that the position of the point in the image after pure rotation depends only on the intrinsic camera parameters, the rotation matrix and the location of the point in the image before the rotation. The 3D point coordinates are not required in the case of pure rotation. As we will see in the nest point, this is not the case when we have rotation with translation.

3.2 The Importance of Pure Rotation

If the axis of rotation does not pass exactly through the optical center of the lens (center of projection) then there will be some translation in addition to the rotation around the center of projection. Considering the existence of a translation vector $T = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T$ the camera coordinate of the point P_c after rotation is obtained using $P'_c = R \cdot P_c + T$ and the location of an image point after rotation and translation will be

$$x'_u = f\frac{r_{11}x_u + r_{12}y_u + r_{13}f + f\frac{t_x}{z_c}}{r_{31}x_u + r_{32}y_u + r_{33}f + f\frac{t_z}{z_c}} \qquad (6)$$

$$y'_u = f\frac{r_{21}x_u + r_{22}y_u + r_{23}f + f\frac{t_y}{z_c}}{r_{31}x_u + r_{32}y_u + r_{33}f + f\frac{t_z}{z_c}}. \qquad (7)$$

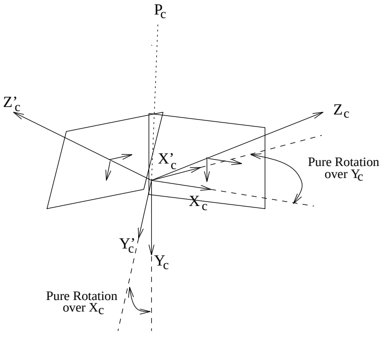

Fig. 5. Pure Rotation around some axis

As we can see from the equation 6 and equation 7 the location of the point in the image after rotation is no longer independent of the depth of the 3D camera point and it also depends on the translation vector. However, if we use feature points that are located far from the camera (z_c considerably large), then the effect of the translation becomes negligible ($z_c \gg ft_x$, $z_c \gg ft_y$, $z_c \gg ft_z$).

3.3 How to Obtain Pure Rotation

As we change the focus or the zoom position of the lens, the optical center of the lens (center of projection) will move along its optical axis. To compensate this displacement the MDOF Active Vision Head build at the ISR-Coimbra has the ability to move the lens along its optical axis with an accuracy of $0.015 \mu m$.

The test whether there is little or no translation is very simple and it involves the use of parallax. The motion of parallax is based on the fact that two 3D points P_{c_1} and P_{c_2} and the center of projection all lie on a straight line, even if we perform pure rotation of the camera along its center of projection. If we don't have pure rotation, and some translation occurs during the rotation, then the three points will no longer be on a straight line and the two points P_{c_1} and P_{c_2} will project at two different image points [1]. Assume that the rotation is around the vertical axis (Y axis of the camera coordinate system). We placed on the wall a pattern with black vertical lines on a white background, printed on a laser printer. To create the parallax effect we placed between the camera and the pattern a transparent acrylic sheet with just one vertical black line. The thickness of this line is less than the thickness of the lines of the pattern in order to create the illusion that this single line is a extension of one of the lines of the pattern. This adjustment has been done by hand, and an edge detector was used to confirm the straightness of the resulting line. If after the rotation the straightness is not the same, this means that we don't have pure rotation and the position of the center of projection must be adjusted by displacing the lens camera body along the optical center adjustment (OCA) degree of freedom.

3.4 The Rotation Method - Implementation and Experimental Details

Since after rotation we have a pair of images, we have chosen to minimize the sum of squared distances between the feature points in the image obtained after rotation and those computed from the initial image and using the pure rotation model described in section III, summed over all the feature points of each pair of images.

To be more precise, the intrinsic parameters can be obtained using N pairs of images taken with the camera rotated at various angles. The relative angles of rotation are measured precisely. The ISR-Coimbra MDOF active vision system has a rotation degree of freedom for each camera (vergence) with a precision of 0.0036°. Corresponding features in each pair of images are found and their pixels coordinates are extracted. There is no special reason to detect the same number M of features in each images, but this is what we did in practice.

We define the cost function:

$$E = \sum_{k=1}^{N} \sum_{n=1}^{M} \left[\left(x^k_{f_{n_2}} - x^k_{f_{n_{rot_{1 \leadsto 2}}}} \right)^2 + \left(y^k_{f_{n_2}} - y^k_{f_{n_{rot_{1 \leadsto 2}}}} \right)^2 \right] \tag{8}$$

where $(x^k_{f_{n_{rot_{i \leadsto j}}}}, y^k_{f_{n_{rot_{i \leadsto j}}}})$ are the coordinates of $(x^k_{f_{n_i}}, y^k_{f_{n_i}})$ after rotation from image i to image j in each pair.

Combining the cost function with equations 4 and 5 and defining the image points on the frame-buffer plane the cost function E can now be defined by

$$E = \sum_{k=1}^{N} \sum_{n=1}^{M} \left[\left(x^k_{f_{n_2}} - x'^k_{f_{n_2}} - c_x \right)^2 + \left(y^k_{f_{n_2}} - y'^k_{f_{n_2}} - c_y \right)^2 \right]$$

where

$$x'^k_{f_{n_2}} = k_x f_x \frac{r_{11}(x^k_{f_{n_1}} - c_x) + r_{12}(y^k_{f_{n_1}} - c_y)k^{-1} + r_{13} f_x k_x}{r_{31}(x^k_{f_{n_1}} - c_x) + r_{32}(y^k_{f_{n_1}} - c_y)k^{-1} + r_{33} f_x k_x}$$

$$y'^k_{f_{n_2}} = k_y f_y \frac{r_{21}(x^k_{f_{n_1}} - c_x)k + r_{22}(y^k_{f_{n_1}} - c_y) + r_{23} f_y k_y}{r_{31}(x^k_{f_{n_1}} - c_x)k + r_{32}(y^k_{f_{n_1}} - c_y) + r_{33} f_y k_y}$$

The task is now to find the intrinsic parameters of the camera ($f_x k_x, f_y k_y, k, c_x, c_y$) by a straightforward nonlinear search, assuming some initial values to the parameters. A detailed description of the approach used to obtain the initial values for the intrinsic parameters can be found on [4].

4. Conclusions

Some characteristics of the ISR MDOF Active Vision Systems and some mathematical background of its optical system calibration have presented in this paper.

This active vision system was designed to be a *Lab* instrument with redundancy on the degrees of freedom and high performances in terms of speeds, acceleration and resolution. Its performances should approach those of a human head. Biological reasons played a role in the design options of this system.

The ability of the active vision systems to perform accurate movements have been used to perform the active calibration using just features of the environment has the camera undergoes an accurate and controllable movement. In the case of the approach used and described on the paper a pure rotation of the camera was considered and a procedure to ensure pure rotation around the center of projection was presented.

ACKNOWLEDGEMENT

This work has been partially supported by Junta Nacional de Investigação Científica e Tecnológica (JNICT) project VARMA.

REFERENCES

[1] A. Yarbus, "Eye Movements and Vision", Plenum Press, New-York, 1967.

[2] A. Basu, "Active Calibration: Alternative Strategy and Analysis", in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 495-500, 1993.

[3] D. Geiger, A. Yuille, "Stereo and Eye Movements" in*Biological Cybernetics*, vol. 62, pp. 117-128, 1989.

---

[1] This observation is not valid for the case of a translation along the projective line defined by the three points. Since the translation is due to rotation about some point, the direction of translation continuously change and therefore only momentarily is aligned with the projective line.

[4] J. Batista, J. Dias, H. Araújo, A. T. Almeida, " The ISR MDOF Active Vision Robot Head - Design and Calibration", SMART Workshop, Lisbon, 1995,

[5] F. Du and M. Brady, "Self-Calibration of the Intrinsic Parameters of Cameras for Active Vision", in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 477-482, 1993.

[6] G. Stein, "Internal Camera Calibration using Rotation and Geometric Shapes", *Msc. Thesis, MIT*, 1993.

[7] H. Christensen, K. Bowyer, H. Bunke, editors, "Active Robot Vision : Camera Heads, model based navigation and reactive control", vol.6, World Scientific, 1994.

[8] J. Crowley, P. Bobet, C. Schmid, "Auto-calibration by direct observation of objects", in *Image and Vision Computing Journal*, vol. 11, no. 2, pp. 67-81, 1993.

[9] J. Lavest, G. Rives, and M. Dhome, "Three-Dimensional Reconstruction by Zooming", in *IEEE Trans. on Robotics and Automation*, vol. 9, No. 2, 1993.

[10] K. Pahlavan, "Active Robot Vision and Primary Ocular Processes", *PhD Thesis, CVAP, KTH*, 1993.

[11] L. Dron, "Dynamic Camera Self-Calibration from Controlled Motion Sequences", in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 501-506, 1993.

[12] M. Li, "Camera Calibration of a Head-Eye System for Active Vision", in *Proc. European Conf. Computer Vision*, pp. 543-554, 1994.

[13] M. Li, "Camera Calibration of the KTH Head-Eye System", in *Technical report, CVAP-147, NADA, KTH*, 1994.

[14] P. MacLauchlan and D. Murray, "Active Camera Calibration for a Head-Eye Platform using a Variable State-Dimension Filter", submitted to *IEEE Trans. on Pattern Analysis and Machine Intelligence*.

[15] R. Carpenter, "Movements of the Eyes", Pion Limited : London, England, 1988.

[16] R. Hartley, "Self-Calibration from Multiple Views with a Rotating Camera", in *Proc. European Conf. Computer Vision*, pp. 471-478, 1994.

[17] R. Willson and S. Shafer, "Precision Imaging and Control for Machine Vision Research at Carnegie Mellon University", in *Proc. SPIE Conf. on High Resolution Sensors and Hybrid Systems*, 1992.

[18] R. Wurtz, M. Goldberg, "The Neurobiology of saccadic Eye Movements", Elsivier, New-York, 1989.

[19] S. Maybank and O. Faugeras, "A Theory of self-calibration of a moving camera", in *International Journal of Computer Vision*, 8:2, pp. 123-151, 1992.

[20] Webb Associates, "Anthropometric Source Book, vol I: Anthropometry for Designers", NASA Reference Publication 1024, 1978.

[21] A. Abbott, N. Ahuja, "Active surface reconstruction by integrating focus, vergence, stereo, and camera calibration", in *3rd Int. Conf. Comp. Vision*, Osaka, Japan, 1990.

[22] T. Poggio et al., " The MIT vision machine", in *Proc. DARPA Image Understanding Workshop*, 1988.

[23] E. Krotkov, F. Fuma, J. Summers, "An agile stereo camera system for flexible image aquisition", in *IEEE J. Robot. Automation*, pp. 108-113, 1988.

[24] C. Brown, "Kinematics and 3D motion predition for gaze control", in *Proc. Workshop on Interpretation of 3D Scenes*, Austin, 1989.

[25] J. Clark, N. Ferrier, "Modal control of an attentive vision system", in *Proc. 2nd Intern. Conf. Comput. Vision*, Tampa, 1988.

[26] J. Dias, J. Batista, C. Simplício, Helder Araújo, A. T. Almeida, "Implementation of an active vision system", in *International Workshop on Mechatronics Computer Systems for Perception and Action*, Halmstad University, Sweden, 1993.

[27] J. Fiala, R, Lumia, K. Roberts, A. Wavering, "TRICLOPS : A tool for studing active vision", in *International Journal of Computer Vision*, 1994.

[28] P. Sharkey, D. Murray, S. Vandevelde, I. Reid, P. Mclauchlan, " A modular head/eye platform for real-time reactive vision" in *International Workshop on Mechatronics Computer Systems for Perception and Action*, Halmstad University, Sweden, 1993.