Proceedings of the 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation, Madrid, Spain, February 15-17, 2006 (pp77-81)

Modelling and calibration of a camera for robot trajectories recording.

V. Niola, C. Rossi, S. Savino
Department of Ingegneria Meccanica per l'Energetica
Università di Napoli - "Federico II"
Via Claudio 21, 80125 Napoli
ITALY

Abstract: - The problem of the camera system modelling is studied and an algorithm for the calibration of the vision system is presented. By means of suitable matrixes a stereoscopic vision system is obtained. This algorithm is suitable to be used to apply vision model to robotic applications

Key words: Vision system, Robot trajectories, Camera Modelling, Camera Calibration

1 Introduction

In many robotic application it is very important to record the robot trajectories; many devices have been developed and used for this purpose. Among them those non invasive seems to be preferable, in fact these last can be easily fitted just when they are necessary and don't need any fixed structure. From this point of view, the device that seems to be preferable are the ones that use vision systems by telecameras [1-3] as it meets completely the requirements mentioned above.

In a previous contribution [4] the authors have proposed a vision algorithm by means of which it can be rather easy to record trajectories of a point belonging to a robot arm in the three dimensional space. This technique can be easily used for many robotic applications and will be particularly useful if it is joined with another algorithm for robot calibration.

The proposed algorithm uses the fourth row of the Denavit and Hartemberg transformation matrix that for kinematics' purposes usually contains three zeros and a scale factor, so it is useful to start from the perspective transform matrix.

In the present contribution the problem of the telecamera system modelling is studied and an algorithm for the stereoscopic vision system is presented. This algorithm is suitable to be used to apply vision model to robotic applications.

2 Prospectic transform and vision systems

A vision system essentially associates a point in the Cartesian space with a point on the image plane. A very common vision system is the television camera that is essentially composed by an optic system (one or more lenses), an image processing and managing system and an image plane; this last is composed by vision sensors. The light from a point in the space is conveyed by the lenses on the image plane and recorded by the vision sensor.

Let us confine ourselves to consider a simple vision system made up by a thin lens and an image plane composed by CCD (*Charged Coupled Device*) sensors. This kind of sensor is a device that is able to record the electric charge that is generated by a photoelectric effect when a photon impacts on the sensor's surface.

It is useful to remember some aspects of the optics in a vision system.

2.1 The thin lenses

A lens is made up by two parts of a spherical surfaces (dioptric surfaces) joined on a same plane. The axis, normal to this plane, is the optical axis. As shown in fig.1, a convergent lens conveys the parallel light rays in a focus F at distance f (focal distance) from the lens plane. The focal distance f, in air, is given by:

$$f = (n-1)\cdot\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

where n is the refractive index of the lens and R_1 ed R_2 are the bending radius of the dioptric surfaces.

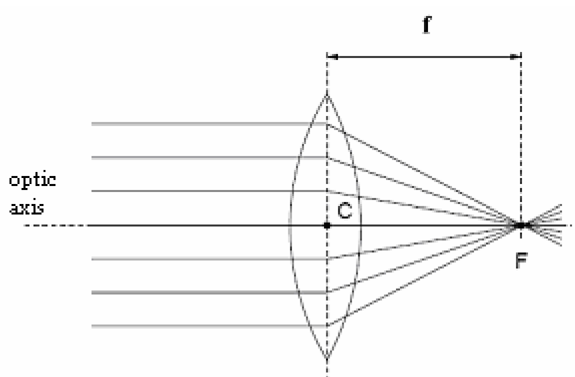

Fig. 1

Now consider a thin lens, a point P and a plane on which the light-rays refracted from the lens are projected as shown in fig.2. the equation for the thin lenses gives:

$$\frac{1}{d}+\frac{1}{L}=\frac{1}{f}$$

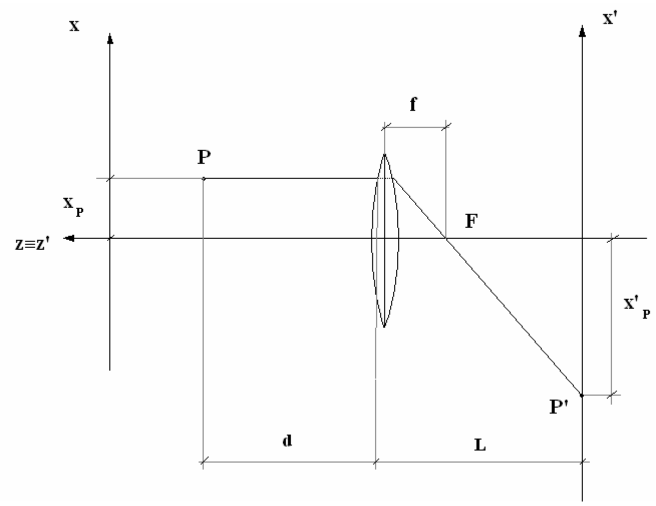

Fig. 2

It is possible to determinate the connection between the position of point P in the space and it's correspondent P' in the projection's plane (see fig.5).
If two frames (xyx for the Cartesian space and x'y'z' for the image plane), having their axes parallel, are assigned and if the thickness of the lens is neglected, from the similitude of the triangles in fig.5 it comes:

$$\frac{x_P}{f}=-\frac{x_P^{'}}{L-f} \qquad (1)$$

From the equation of the thin lenses:

$$L=\frac{f\cdot d}{d-f} \quad \Rightarrow L-f=\frac{f^2}{d-f}$$

Hence:

$$x_P^{'}=-\frac{f}{d-f}\cdot x_P$$

If we consider that generally the distance of a point from the camera's objective is one meter or more while the focal distance is about some millimetres (*d*>>*f*), the following approximation can be accepted:

$$x_P^{'}\cong-\frac{f}{d}\cdot x_p$$

So the coordinates of the point in the image plane can be obtained by scaling the coordinates in the Cartesian space by a factor $-d/f$. The minus sign is due to the upsetting of the image.

3 The model of the telecamera

As already observed a telecamera can be modelled as a thin lens and an image plane with CCD sensors. The objects located in the Cartesian space emit rays of light that are refracted from the lens on the image plane. Each CCD sensor emit an electric signal that is proportional to the intensity of the ray of light on it; the image is made up by a number of pixels, each one of them records the information coming from the sensor that corresponds to that pixel.
In order to indicate the position of a point of an image it is possible to define a frame u,v (see fig.3) which axes are contained in the image plane. To a given point in the space (which position is given by its Cartesian coordinates) it is possible to associate a point in the image plane (two coordinates) by means of the telecamera. So, the expression "model of the camera" means the transform that associates a point in the Cartesian space to a point in the image space.
It has to be said that in the Cartesian space a point position is given by three coordinates expressed in length unit while in the image plane the two coordinates are expressed in pixel; this last is the smaller length unit that ca be revealed by the camera and isn't a normalized length unit. The model of the camera must take onto account this aspect also.
In order to obtain the model of the camera the scheme reported in fig.3 can be considered.

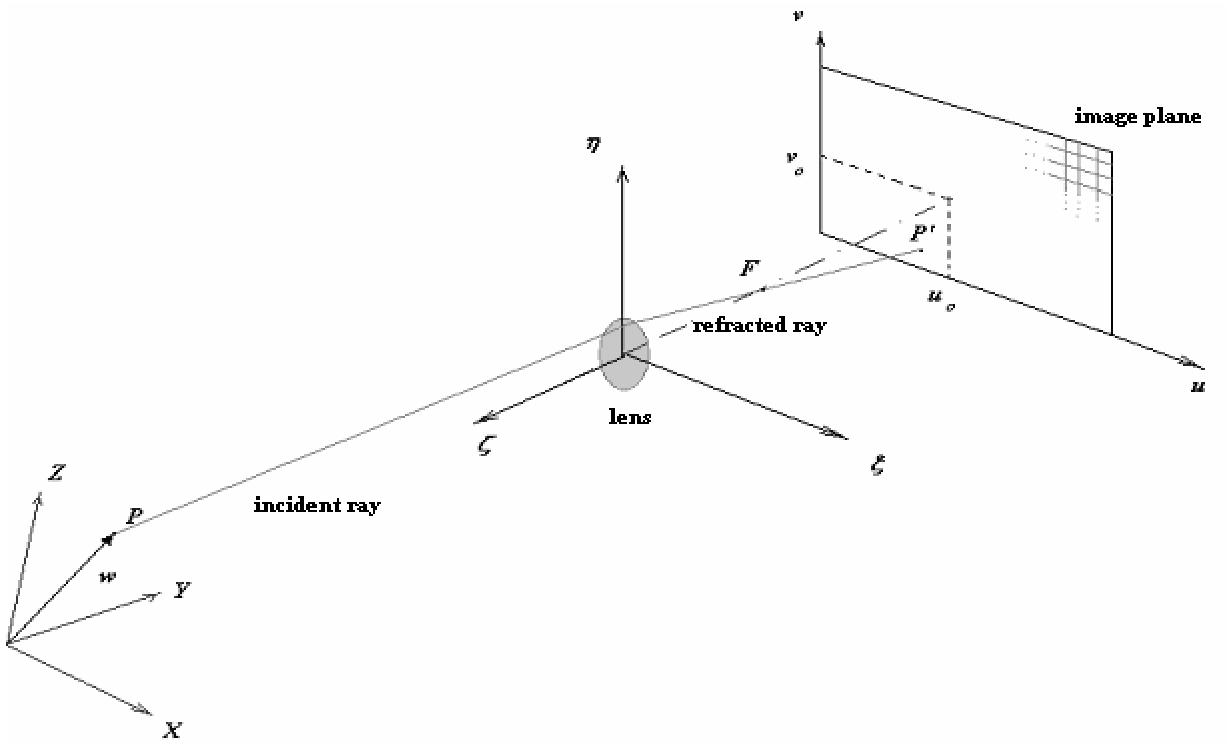

Fig.3

Consider a frame xyz in the Caresian space, the position of a generic point P in the space is given by the vector w. Then consider a frame ξηζ having the origin in the lens centre and the plane ξη coincident with the plane of the lens; hence, the plane ξη is parallel to the image plane and ζ axis is coincident with the optical axis. Finally consider a frame u,v on the image plane so that u_o and v_o are the coordinates of the origin of frame ξηζ, expressed in pixel.

As it was already told, the lens makes a perspective transform in which the constant of proportionality is –f . If this transform is applied to vector w, a w_l vector is obtained:

$$\tilde{w}_l = T_l \cdot \tilde{w} \qquad (2)$$

Were the matrix T_l is obtained dividing by –f the last row of the perspective transformation matrix T_p, that is described in[2]:

$$T_l = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ -\frac{D_x}{f} & -\frac{D_y}{f} & -\frac{D_z}{f} & 0 \end{bmatrix} \qquad (3)$$

Substantially, the above essentially consists in a changing of the reference frames and a scaling based on the rules of geometric optics previously reported.

Assumed x_l e y_l as the first two components of the vector w_l, the coordinates u and v (expressed in pixel) of P' (image of P) are :

$$\begin{cases} u = \frac{x_l}{\delta_u} + u_o \\ v = \frac{x_l}{\delta_v} + v_o \end{cases} \qquad (4)$$

Where δ_u e δ_v are respectively the horizontal and vertical dimensions of the pixel.

So, by substituting eq,(2) in eq.(4) it comes:

$$\begin{cases} u = -\frac{f}{D^T w}\left[\left(\frac{1}{\delta_u}\cdot\hat{\xi} - \frac{u_o}{f}\cdot D\right)^T w + \frac{1}{\delta_u}\cdot t_\xi\right] \\ v = -\frac{f}{D^T w}\left[\left(\frac{1}{\delta_v}\cdot\hat{\eta} - \frac{v_o}{f}\cdot D\right)^T w + \frac{1}{\delta_v}\cdot t_\eta\right] \end{cases}$$

(5)

Finally if we define the vector $m = [u \quad v]^T$, the representation in homogeneous coordinates $\tilde{m} = [m_1 \quad m_2 \quad -D^T w/f]^T = [u \quad v \quad -D^T w/f]^T$ of the previous vector can be written :

$$\tilde{m} = M \cdot \tilde{w} \qquad (6)$$

Where M is the matrix :

$$M = \begin{bmatrix} \left(\frac{\xi_x}{\delta_u} - \frac{u_o D_x}{f}\right) & \left(\frac{\xi_y}{\delta_u} - \frac{u_o D_y}{f}\right) & \left(\frac{\xi_z}{\delta_u} - \frac{u_o D_z}{f}\right) & t_\xi / \delta_u \\ \left(\frac{\eta_x}{\delta_v} - \frac{v_o D_x}{f}\right) & \left(\frac{\eta_y}{\delta_v} - \frac{v_o D_y}{f}\right) & \left(\frac{\eta_z}{\delta_v} - \frac{v_o D_z}{f}\right) & t_\eta / \delta_v \\ -D_x / f & -D_y / f & -D_z / f & 0 \end{bmatrix} \quad (7)$$

that represents the requested model of the camera.

4 The stereoscopic vision

That above reported concurs to determine the coordinates in image plane (u,v) of a generic point of tridimensional space $w=(w_x\ w_y\ w_z)^T$, but the situation is more complex if it is necessary to recognise the position (w) of a point starting to its camera image (u, v). In this case the expression (5) becomes a system of 2 equation with 3 unknowns, so it isn't absolutely solvable.
This obstacle can be exceeded by means of a vision system with at least two cameras.
In this way, that above reported can be applied to the recording of a robot trajectory in the tridimensional space by using two cameras. This will emulate the human vision.
Let us consider two cameras and say M and M' their transform matrixes. We want to recognise the position of a point P, that in the Cartesian space is given by a vector w in a generic frame xyz. From eq.(6) we have:

$$\begin{cases} \tilde{m} = M \cdot w \\ \tilde{m}' = M' \cdot w \end{cases} \quad (8)$$

where:

$$\tilde{m} = \begin{pmatrix} m_1 \\ m_2 \\ -D^T w / f \end{pmatrix}$$

and

$$\tilde{m}' = \begin{pmatrix} m_1' \\ m_2' \\ -D'^T w / f' \end{pmatrix}$$

are the position vectors in the image plane of the cameras, in homogeneous coordinates, and:

$$M = \begin{bmatrix} \mu_1^T & \mu_{14} \\ \mu_2^T & \mu_{24} \\ -D^T / f & 0 \end{bmatrix};$$

(9)

$$M' = \begin{bmatrix} \mu'^T_1 & \mu'_{14} \\ \mu'^T_2 & \mu'_{24} \\ -D'^T / f' & 0 \end{bmatrix}$$

are the [3x4] transformation matrixes (7) from spatial frame to image planes of two cameras.
The first equation of the system (8), in Cartesian coordinates (non-homogenous), can be written as:

$$\begin{cases} u = -f \cdot \dfrac{\mu_1^T w + \mu_{14}}{n^T w} \\ v = -f \cdot \dfrac{\mu_2^T w + \mu_{24}}{n^T w} \end{cases} \quad (10)$$

or :

$$\begin{cases} (u \cdot D + f \cdot \mu_1)^T w = \mu_{14} \\ (v \cdot D + f \cdot \mu_2)^T w = \mu_{24} \end{cases} \quad (11)$$

In the same way for the camera whose transform matrix is M', it can be written:

$$\begin{cases} (u' \cdot D' + f' \cdot \mu'_1)^T w = \mu'_{14} \\ (v' \cdot D' + f' \cdot \mu'_2)^T w = \mu'_{24} \end{cases} \quad (12)$$

By arranging eq.(33) and eq.(34) we obtain:

$$\begin{bmatrix} (u \cdot D + f \cdot \mu_1)^T \\ (v \cdot D + f \cdot \mu_2)^T \\ (u' \cdot D' + f' \cdot \mu'_1)^T \\ (v' \cdot D' + f' \cdot \mu'_2)^T \end{bmatrix} \cdot w = \begin{bmatrix} \mu_{14} \\ \mu_{24} \\ \mu'_{14} \\ \mu'_{24} \end{bmatrix} \quad (13)$$

This last equation represents the stereoscopic problem and consist in a system of 4 equation in 3 unknown (w_x, w_y, w_z). As the equations are more than the unknowns can be solved by a least square algorithm. In this way it is possible to invert the problem that is described by eqs.(5) and to recognise the position of a generic point starting to its camera image.

5 Conclusions

By means of studies on a camera vision model, an algorithm for stereoscopic vision system is presented.

The proposal algorithm has been not jet experimentally tested, it is a method in order to use a vision model on robotic application.

Some tests are now in progress to verify the reliability of algorithm in relation with different calibration procedures of camera system. The results of these tests will be used to introduce this algorithm in robot trajectories acquisition and represent first step to implement a robot vision control technique.

References

[1] A. FUSIELLO – "Visione Computazionale: appunti delle lezioni" – Informatic Department, University of Verona, 3 March 2005.

[2] V. Niola, C. Rossi, S. Savino – "Perspective Transform and Vision System for Robot Trajectories Recording" - Proc. 5th WSEAS Int. Conf. on "Signal Processing, Robotics and Automation" – Madrid, February 15-17, 2006.

[3] R. BRANCATI, C. ROSSI, S. SCOCCA - Trajectories Recording for Investigations on Robot Dynamics and Mechanical Calibrations – Proc. of 10th int. Workshop RAAD 2001, Wien, May 16-19, 2001

[4] NIOLA, V. ROSSI C.. - "A method for the trajectories recording of a robot arm: early results - Proc. 9th WSEAS International Conference on Computers, Atene 11 – 16 July 2005.

[5] NIOLA V., ROSSI C. – "Video acquisition of a robot arm trajectories in the work space" - WSEAS Transactions on Computers, Iusse 7, Vol. 4, july 2005, pagg. 830 – 836.

[6] NASTI G., NIOLA V., SAVINO S. – "Experimental Study on Vision Motion Measures in Robotics" - 9*th* WSEAS International Conference on Computers, Vouliagmeni – Athens, Greece, July 11 – 16, 2005.

[7] R. SHARMA, S, HUTCHINSON - Motion perceptibility and its application to active vision-based servo control - Technical Report UIUC-BI AI RCV-94-05, The Beckman Institute, Università dell'Illinois, 1994.

[8] R. SHARMA - Active vision for visual servong: a review - IEEE Workshop on Visual Servoing: Achivement, Application and Open Problems, Maggio 1994.

[9] R. SHARMA, S, HUTCHINSON - Optimizing hand/eye configuration for visual-servo system - IEEE International Conference on Robotics and Automation , pp. 172-177, 1995.