We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

14

Video Applications for Robot Arms

Vincenzo Niola, Cesare Rossi
Department of Ingegneria Meccanica per l'Energetica
University of Napoli - "Federico II"
Via Claudio 21, 80125 Napoli
ITALY

1. Introduction

The "Artificial Vision" permits industrial automation and system vision able to act in the production activities without humane presence. So we suppose that the acquisition and interpretation of the imagines for automation purposes is an interesting topic.

Industrial application are referred to technological fields (assembly or dismounting, cut or stock removal; electrochemical processes; abrasive trials; cold or warm moulding; design with CAD techniques; metrology), or about several processes (control of the row material; workmanship of the component; assemblage; packing or storages; controls of quality; maintenance).

The main advantages of this technique are:

1) elimination of the human errors, particularly in the case of repetitive or monotonous operations;
2) possibility to vary the production acting on the power of the automatic system (the automatic machines can operate to high rhythms day and night every day of the year);
3) greater informative control through the acquisition of historical data; these data can be used for successive elaborations, for the analysis of the failures and to have statistics in real time;
4) quality control founded on objective parameters in order to avoid dispute, and loss of image.

In addition, another interesting application can be considered: the robot cinematic calibration and the trajectories recording. First of all it is important to consider that, by a suitable cameras calibration technique, it is possible to record three dimensional objects and trajectories by means of a couple of television cameras. By means of perspective transformation it is possible to associate a point in the geometric space to a point in a plane. In homomogeneous coordinates the perspective transformation matrix has non-zero elements in the fourth row.

2. The perspective transform

In this paragraph, an expression of perspective transformation is proposed, in order to introduce the perspective concepts for the application in robotic field.

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.

2.1 The matrix for the perspective transformation

It is useful to remember that by means of a perspective transform it is possible to associate a point in the geometric space to a point in a plane, that will be called "image plane"; this will be made by using a scale factor that depends on the distance between the point itself and the image plane.

Let's consider fig.1: the position of point P in the frame O,x,y,z is given by the vector $\mathbf{w}$, while the same position in the frame Ω,ξ,η,ζ is given by vector $\mathbf{w_r}$ and the image plane is indicated with $\mathcal{R}$ this last, for the sake of simplicity is supposed to be coincident with the plane ξ,η.
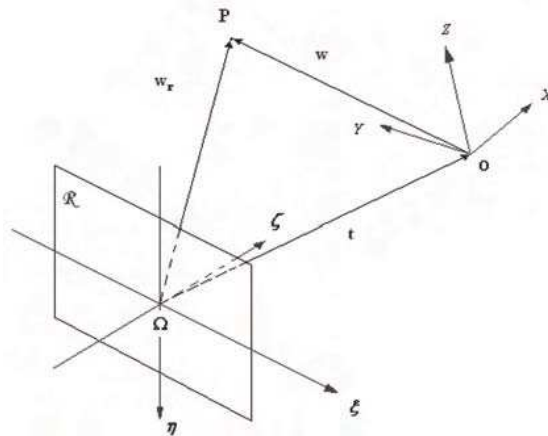

Fig. 1. Frames for the perspective transformation.

The vectors above are joined by the equation:

$$\begin{Bmatrix} w_{r,x} \\ w_{r,y} \\ w_{r,z} \\ sf \end{Bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_\xi \\ R_{21} & R_{22} & R_{23} & t_\eta \\ R_{31} & R_{32} & R_{33} & t_\zeta \\ 0 & 0 & 0 & sf \end{bmatrix} = \begin{Bmatrix} w_x \\ w_y \\ w_z \\ sf \end{Bmatrix} \tag{1}$$

where sf is the scale factor; more concisely equation (1) can be written as follows:

$$\widetilde{w}_r = T \cdot \widetilde{w} \tag{2}$$

where the tilde indicates that the vectors are expressed in homogeneous coordinates. The matrix T is a generic transformation matrix that is structured according to the following template:

Rotation Matrix	Position Vector
Perspective	Scale

The scale factor will almost always be 1 and the perspective part will be all zeros except when modelling cameras.

The fourth row of matrix [T] contains three zeros; as for these last by means of the prospectic transform three values, generally different by zero, will be determined.

Lets consider, now, fig. 2: the vector **w***, that represents the projection of vector **wr** on the plane ξ,η.
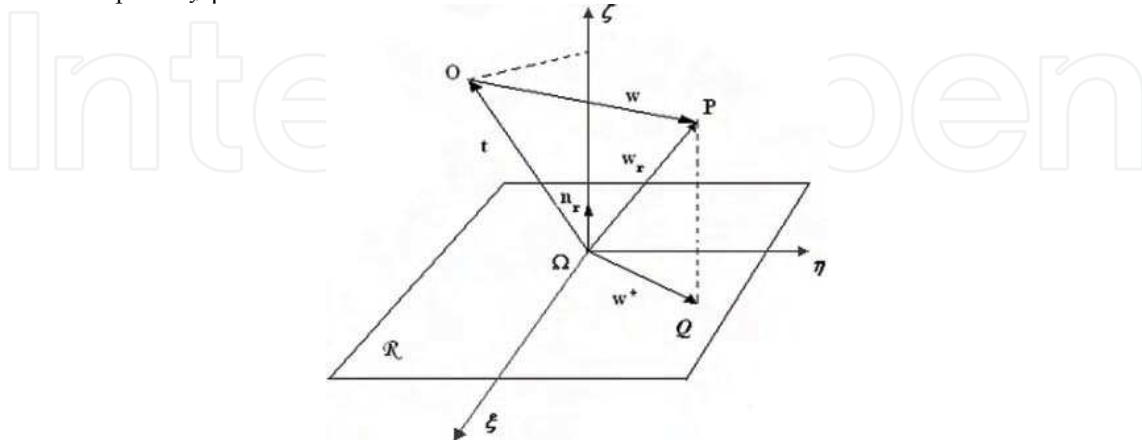

Fig. 2. Vectors for the perspective transformation.

The coordinates of point P in the image plane can be obtained from the vector **w_r**, in fact, these coordinates are the coordinates of **w***, that can be obtained as follows:
Let's consider the matrix R:

$$R = \begin{bmatrix} \hat{\xi}^T \\ \hat{\eta}^T \\ \hat{\zeta}^T \end{bmatrix} \tag{3}$$

where $\hat{\xi}$ $\hat{\eta}$ $\hat{\zeta}$ are the versor of the frame {Ω,ξ,η,ζ} axes in the frame {O,x,y,z}.
In fig.2 the vector **t** indicates the origin of frame O,x,y,z in the frame Ω,ξ,η,ζ and the projection of P on the plane ξ,η is represented by point Q, which position vector is **w***. This last, in homogeneous coordinates is given by:

$$\widetilde{w}^* = \begin{pmatrix} w_{r,\xi} \\ w_{r,\eta} \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \hat{\xi}^T w + t_\xi \\ \hat{\eta}^T w + t_\eta \\ 0 \\ 1 \end{pmatrix} \tag{4}$$

In the same figure, **n_r** is the versor normal to the image plane R , and **n** will be the same versor in the frame {O,X,Y,Z}. The perspective image of vector **w*** can be obtained by assessing a suitable scale factor. This last depends on the distance d between point P and the image plane. The distance d is given from the following scalar product:

$$d = n_r{}^T w_r \tag{5}$$

Let's indicate with **w**{Ω,ξ,η,ζ} the vector **w** in the frame {Ω,ξ,η,ζ}:

$$\widetilde{w}_{\{\Omega,\xi,\eta,\zeta\}} = \begin{pmatrix} w_\xi \\ w_\eta \\ w_\zeta \\ 1 \end{pmatrix}$$

Because $\hat{\xi}$ $\hat{\eta}$ $\hat{\zeta}$ are the versor of the frame $\{\Omega,\xi,\eta,\zeta\}$ axes in the frame $\{O,x,y,z\}$, it is possible to write the coordinates of the vector $\mathbf{w}\{\Omega,\xi,\eta,\zeta\}$ in the frame $\{\Omega,\xi,\eta,\zeta\}$:

$$w_\xi = \hat{\xi}^T \cdot w = \xi_x w_x + \xi_y w_y + \xi_z w_z;$$
$$w_\eta = \hat{\eta}^T \cdot w = \eta_x w_x + \eta_y w_y + \eta_z w_z;$$
$$w_\zeta = \hat{\zeta}^T \cdot w = \zeta_x w_x + \zeta_y w_y + \zeta_z w_z.$$

In the frame $\{\Omega,\xi,\eta,\zeta\}$, it is possibile to write $\mathbf{w_r}$ as sum of $\mathbf{w}\{\Omega,\xi,\eta,\zeta\}$ and $\mathbf{t}$:

$$\widetilde{w}_r = \widetilde{w}_{\{\Omega,\xi,\eta,\zeta\}} + \widetilde{t} = \begin{pmatrix} w_\xi + t_\xi \\ w_\eta + t_\eta \\ w_\zeta + t_\zeta \\ 1 \end{pmatrix} = $$
$$= \begin{pmatrix} \xi_x w_x + \xi_y w_y + \xi_z w_z + t_\xi \\ \eta_x w_x + \eta_y w_y + \eta_z w_z + t_\eta \\ \zeta_x w_x + \zeta_y w_y + \zeta_z w_z + t_\zeta \\ 1 \end{pmatrix}$$

(5′)

an expression of d is:

$$d = n_r^{\ T} w_r = \begin{pmatrix} n_{r,\xi} \\ n_{r,\eta} \\ n_{r,\zeta} \end{pmatrix}^T \cdot \begin{pmatrix} w_\xi + t_\xi \\ w_\eta + t_\eta \\ w_\zeta + t_\zeta \end{pmatrix} = $$
$$= \begin{pmatrix} n_{r,\xi} \\ n_{r,\eta} \\ n_{r,\zeta} \end{pmatrix}^T \cdot \begin{pmatrix} \xi_x w_x + \xi_y w_y + \xi_z w_z + t_\xi \\ \eta_x w_x + \eta_y w_y + \eta_z w_z + t_\eta \\ \zeta_x w_x + \zeta_y w_y + \zeta_z w_z + t_\zeta \end{pmatrix}$$

Let's introduce the expressions:

$$D_x = \frac{(\xi_x w_x + \xi_y w_y + \xi_z w_z + t_\xi) \cdot n_{r,\xi}}{w_x};$$
$$D_y = \frac{(\eta_x w_x + \eta_y w_y + \eta_z w_z + t_\eta) \cdot n_{r,\eta}}{w_y};$$
$$D_z = \frac{(\zeta_x w_x + \zeta_y w_y + \zeta_z w_z + t_\zeta) \cdot n_{r,\zeta}}{w_z};$$

it is possible to write:

$$d = n_r^{\ T} w_r = \begin{pmatrix} D_x \\ D_y \\ D_z \\ 0 \end{pmatrix}^T \cdot \begin{pmatrix} w_x \\ w_y \\ w_z \\ 1 \end{pmatrix} = D^T \cdot w$$

(5″)

In the expression (5″) the vector D is:

$$D = \begin{pmatrix} D_x \\ D_y \\ D_z \\ 0 \end{pmatrix}$$

As vector **w*** is given by:

$$\widetilde{w}^*{}_p = \begin{pmatrix} \hat{\xi}^T w + t_\xi \\ \hat{\eta}^T w + t_\eta \\ 0 \\ n_r{}^T w_r \end{pmatrix} \tag{6}$$

The perspective matrix [Tp] can be obtained:

$$\widetilde{w}^*{}_p = T_p \cdot \widetilde{w} \qquad \Rightarrow \qquad T_p = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ D_x & D_y & D_z & 0 \end{bmatrix} \tag{7}$$

The terms Dx, Dy, Dz assume infinity values if the vector **w** has one of his coordinates null, but this does not influence on generality of the relation $\widetilde{w}^*{}_p = T_p \cdot \widetilde{w}$, in fact in p this case, the term that assume infinity value, is multiplied for zero.

2.2 The perspective concept

From equation 7 some useful properties can be obtained in order to define how a geometric locus changes its representation when a perspective transform occurs. As for an example of the above said, let us consider the representation of the displacement of a point in the space: suppose that the displacement occurs, initially, in the positive direction of x axis. Say this displacement Δ w, the point moves from the position P to the position P′, that are given by the vectors:

$$w = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} \quad e \quad w' = \begin{pmatrix} w'_x \\ w_y \\ w_z \end{pmatrix} \tag{8}$$

If the perspective transforms are applied we have :

$$p = T_p \cdot w \qquad \text{and} \quad p' = T_p \cdot w'$$

so:

$$p = \begin{pmatrix} \dfrac{\xi_x w_x + \xi_y w_y + \xi_z w_z + t_\xi}{D^T w} \\ \dfrac{\eta_x w_x + \eta_y w_y + \eta_z w_z + t_\eta}{D^T w} \\ 0 \\ \end{pmatrix}$$

and

$$p' = \begin{pmatrix} \dfrac{\xi_x w'_x + \xi_y w_y + \xi_z w_z + t_\xi}{D'^T w'} \\ \dfrac{\eta_x w'_x + \eta_y w_y + \eta_z w_z + t_\eta}{D'^T w'} \\ 0 \end{pmatrix} \qquad (9)$$

Hence, the displacement in the image plane is given by:

$$\Delta p \quad = p' - p$$

that is to say:

$$\Delta p = \left\{ \begin{array}{c} \dfrac{\xi_x \cdot [w_x'(D^T w) - w_x(D'^T w')]}{(D^T w)(D'^T w')} \\ \dfrac{\eta_x \cdot [w_x'(D^T w) - w_x(D'^T w')]}{(D^T w)(D'^T w')} \\ 0 \end{array} \right\} \qquad (10)$$

In this way, a displacement Δw along the x axis corresponds to a displacement Δp in the image plane along a straight line which pitch is η_x/ξ_x. So the x axis equation in the image plane is :

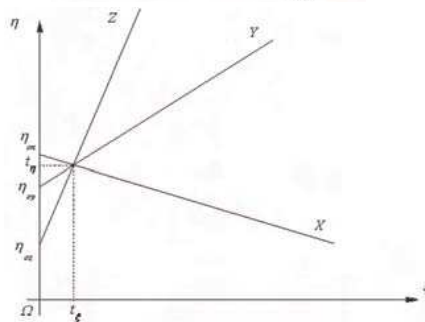$$\eta = (\eta_x/\xi_x) \cdot \xi + \frac{\xi_x t_\eta - \eta_x t_\xi}{\xi_x} \qquad (11)$$

Fig. 3. Frames.

The interception was calculated by imposing that the point which cordinates are belongs to the x axis. In the same way it is possible to obtain the y axis and the z axis equations:

$$\text{y axis}: \quad \eta = (\eta_y/\xi_y) \cdot \xi + \frac{\xi_y t_\eta - \eta_y t_\xi}{\xi_y} \qquad (12)$$

$$\text{z axis}: \quad \eta = (\eta_z/\xi_z) \cdot \xi + \frac{\xi_z t_\eta - \eta_z t_\xi}{\xi_z} \qquad (13)$$

By means of equations (11), (12) and (13) it is possible to obtain a perspective representation of a frame belonging to the Cartesian space in the image plane; that is to say: for a given body it is possible to define it's orientation (e.g. roll, pitch and yaw) in the image plane.

An example could clarify what exposed above: let's consider a circumference which equation in the frame xyz is:

$$x^2 + y^2 = \rho^2 \tag{14}$$

It is possible to associate the geometric locus described from equation (14) to the corresponding one in the image plane; in fact by means of equations (11),(12) e (13) it is possible to write:

$$x = [\xi, \frac{\eta_x}{\xi_x} \cdot \xi + \eta_{0x}]$$
$$y = [\frac{\xi_y}{\eta_y} \cdot (\eta - \eta_{0y}), \eta] \tag{15}$$

By substituting these last in the (14) we obtain :

$$\left[1 + \left(\frac{\eta_x}{\xi_x}\right)^2\right] \cdot \xi^2 + \left[1 + \left(\frac{\xi_y}{\eta_y}\right)^2\right] \cdot \eta^2 + \left(\frac{2\eta_x\eta_{0x}}{\xi_x}\right) \cdot$$
$$\cdot \xi + \left(\frac{2\eta_{0y}\xi_y^2}{\eta_y^2}\right) \cdot \eta = \rho^2 - \left(\frac{\eta_{0x}^2\eta_y^2 - \eta_{0y}^2\xi_y^2}{\eta_y^2}\right)$$

That represents the equation of a conic section. In particular a circumference the centre of which is in the origin of the xyz frame that becomes an ellipse having its foci on a generic straight line in the image plane.

2.3 Perspective transformation in D-H robotic matrix

For kinematics' purposes in robotic applications, it is possible to use the Denavit and Hartemberg transformation matrix in homogeneous coordinates in order to characterize the end-effector position in the robot base frame by means of joints variable, this matrix usually contains three zeros and a scale factor in the fourth row. The general expression of the homogenous transformation matrix that allows to transform the coordinates from the frame i to frame i-1, is:

$$A_i^{i-1} = \begin{bmatrix} C\vartheta_i & -C\alpha_i \cdot S\vartheta_i & S\alpha_i \cdot S\vartheta_i & a_i \cdot C\vartheta_i \\ S\vartheta_i & C\alpha_i \cdot C\vartheta_i & -S\alpha_i \cdot C\vartheta_i & a_i \cdot S\vartheta_i \\ 0 & S\alpha_i & C\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

For a generic robot with n d.o.f., the transformation matrix from end-effector frame to base frame, has the following expression:

$$T_n^0 = A_1^0 \cdot A_2^1 \cdot A_3^2 \cdot \dots \cdot A_n^{n-1}$$

With this matrix it is possible to solve the expression:

$$\{P\}_0 = T_n^0 \cdot \{P\}_n$$

where $\{P\}_0$ and $\{P\}_n$ are the vectors that represent a generic point P in frame 0 and P n frame n.
Can be useful to include the perspective concepts in this trasformation matrix, in this way it is possible to obtain a perspective representation of the robot base frame, belonging to the Cartesian space, in an image plane, like following expression shows:

$$\{P\}_p = T_p \cdot \{P\}_0 = T_p \cdot T_n^0 \cdot \{P\}_n = [T_p]_n^0 \cdot \{P\}_n \tag{16}$$

where $\{P\}_p$ is the perspective image of generic point P and $[T_p]_n^0$ is the perspective transformation matrix from end-effector frame to an image plane.
With this representation the fourth row of the Denavit and Hartemberg matrix will contain non-zero elements. A vision system demands an application like this.

3. The camera model

When vision systems are used for robotic applications, it is important to have a suitable model of the cameras.

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.

3.1 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.4, a convergent lens conveys the parallel light rays in a focus F at distance f (focal distance) from the lens plane.
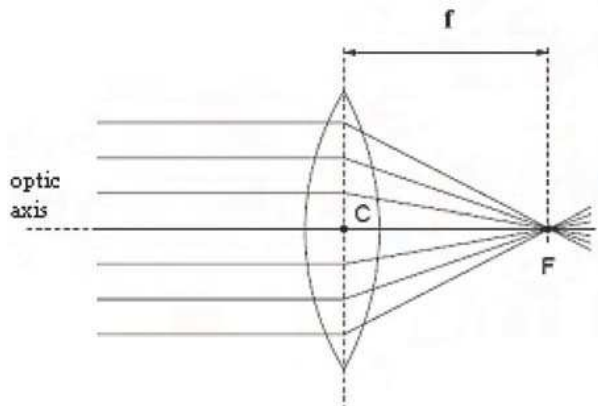

Fig. 4. Convergent lens.

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 *R1* ed *R2* are the bending radius of the dioptric surfaces.

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. 5. the equation for the thin lenses gives:
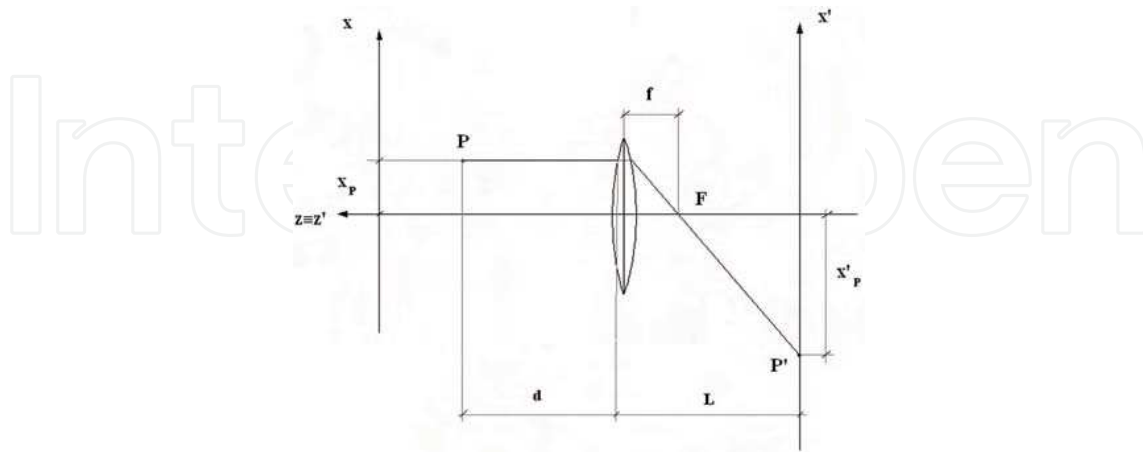
$$\frac{1}{d} + \frac{1}{L} = \frac{1}{f}$$

Fig. 5. Thin lens.

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 (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:

$$\frac{x_P^{'}}{f} = -\frac{x_P}{L-f}$$

From the equation of the thin lenses:

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

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 ($d \gg f$), 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.

3.2 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 (fig. 6) 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.6 can be considered.

Consider a frame xyz in the Cartesian space, the position of a generic point P in the space is given by the vector w. Then consider a frame $\xi\eta\zeta$ having the origin in the lens centre and the plane $\zeta\eta$ coincident with the plane of the lens; hence, the plane $\zeta\eta$ 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 uo and vo are the coordinates of the origin of frame $\xi\eta\zeta$, expressed in pixel.
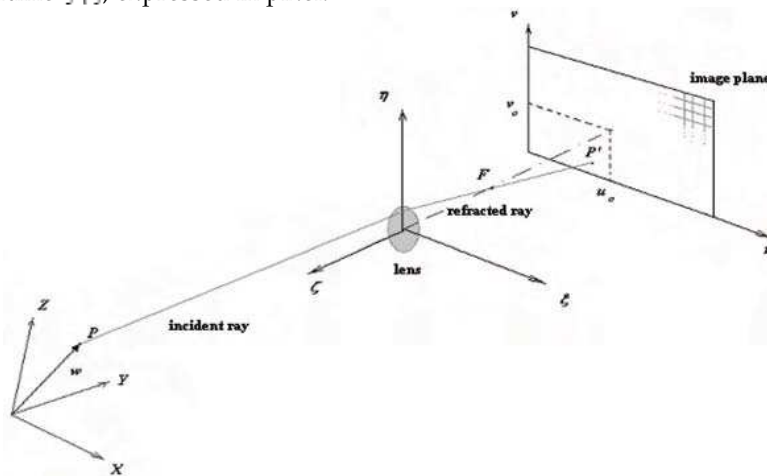

Fig.6. Camera model.

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:

$$\widetilde{w}_l = T_l \cdot \widetilde{w} \tag{17}$$

Were the matrix T_l is obtained dividing by –f the last row of the perspective transformation matrix Tp, (7).

$$T_l = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ -\dfrac{D_x}{f} & -\dfrac{D_y}{f} & -\dfrac{D_z}{f} & 0 \end{bmatrix} \tag{18}$$

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 wl, the coordinates u and v (expressed in pixel) of P' (image of P) are :

$$\begin{cases} u = \dfrac{x_1}{\delta_u} + u_o \\[2mm] v = \dfrac{x_1}{\delta_v} + v_o \end{cases} \tag{19}$$

Where δ_u e δ_v are respectively the horizontal and vertical dimensions of the pixel. So, by substituting equation (17) in equation (19) it comes:

$$\begin{cases} u = -\dfrac{f}{D^T w}\left[\left(\dfrac{1}{\delta_u}\cdot\hat{\xi} - \dfrac{u_o}{f}\cdot D\right)^T w + \dfrac{1}{\delta_u}\cdot t_\xi\right] \\[3mm] v = -\dfrac{f}{D^T w}\left[\left(\dfrac{1}{\delta_v}\cdot\hat{\eta} - \dfrac{v_o}{f}\cdot D\right)^T w + \dfrac{1}{\delta_v}\cdot t_\eta\right] \end{cases} \tag{20}$$

Finally if we define the vector $m = [u \quad v]^T$, the representation in homogeneous $w \ T \ w$ $T\sim$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 :

$$\tilde{m} = M \cdot \tilde{w} \tag{21}$$

Where M is the matrix :

$$M = \begin{bmatrix} \left(\dfrac{\xi_x}{\delta_u} - \dfrac{u_o D_x}{f}\right) & \left(\dfrac{\xi_y}{\delta_u} - \dfrac{u_o D_y}{f}\right) & \left(\dfrac{\xi_z}{\delta_u} - \dfrac{u_o D_z}{f}\right) & t_\xi/\delta_u \\[3mm] \left(\dfrac{\eta_x}{\delta_v} - \dfrac{v_o D_x}{f}\right) & \left(\dfrac{\eta_y}{\delta_v} - \dfrac{v_o D_y}{f}\right) & \left(\dfrac{\eta_z}{\delta_v} - \dfrac{v_o D_z}{f}\right) & t_\eta/\delta_v \\[3mm] -D_x/f & -D_y/f & -D_z/f & 0 \end{bmatrix} \tag{22}$$

that represents the requested model of the camera.

3.3 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 1)^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 (20) 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 three dimensional 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 equation (21) we have:

$$\begin{cases} \tilde{m} = M \cdot w \\ \tilde{m}' = M' \cdot w \end{cases} \tag{23}$$

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:

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

(24)

$$M' = \begin{bmatrix} \mu_1'^T & \mu_{14}' \\ \mu_2'^T & \mu_{24}' \\ -D'^T/f' & 0 \end{bmatrix}$$

are the [3x4] transformation matrixes (22) from spatial frame to image planes of two cameras.
The first equation of the system (23), in Cartesian coordinates (non-homogenous), can be written as:

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

(25)

or:

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

(26)

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

$$\begin{cases} (u' \cdot D' + f' \cdot \mu_1')^T w = \mu_{14}' \\ (v' \cdot D' + f' \cdot \mu_2')^T w = \mu_{24}' \end{cases}$$

(27)

By arranging eq.(26) and eq.(27) we obtain:

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

(28)

This last equation represents the stereoscopic problem and consist in a system of 4 equation in 3 unknown (*wx,wy,wz*). 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 equations (20) and to recognise the position of a generic point starting to its camera image.

3.4 The stereoscopic problem

Relation (28) represents the stereoscopic problem, it consists in a system of 4 equations in 3 unknown, in the form:

$$A(u, u', v', w) \cdot w = B$$

where A is a matrix that depends by two couple of camera coordinates (u,v) and (u',v'), and by vector w, and B is a vector with parameters of cameras configuration.
It is possible to find an explicit form of this problem.

Vectors D and w have the following expression:

$$w = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix};$$

$$D = \begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix} = \begin{pmatrix} \dfrac{(\xi_x w_x + \xi_y w_y + \xi_z w_z + t_\xi) \cdot n_{r,\xi}}{w_x} \\ \dfrac{(\eta_x w_x + \eta_y w_y + \eta_z w_z + t_\eta) \cdot n_{r,\eta}}{w_y} \\ \dfrac{(\zeta_x w_x + \zeta_y w_y + \zeta_z w_z + t_\varsigma) \cdot n_{r,\varsigma}}{w_z} \end{pmatrix}$$

(29)

Starting to first equation of (20), it is possible to write:

$$u = -\frac{f}{D^T w} \left[\left(\frac{1}{\delta_u} \cdot \hat{\xi} - \frac{u_0}{f} \cdot D \right)^T w + \frac{1}{\delta_u} \cdot t_\xi \right] \Rightarrow$$

$$\left(\frac{\xi_x}{\delta_u} - \frac{u_0 \cdot D_x}{f} \right) \cdot w_x + \left(\frac{\eta_x}{\delta_u} - \frac{u_0 \cdot D_y}{f} \right) \cdot w_y + \left(\frac{\zeta_x}{\delta_u} - \frac{u_0 \cdot D_z}{f} \right) \cdot w_z + \frac{t_\xi}{\delta_u} =$$

$$-\frac{u}{f} \cdot \left[D_x \cdot w_x + D_y \cdot w_y + D_z \cdot w_z \right] \Rightarrow$$

$$\frac{1}{\delta_u} \left(\xi_x \cdot w_x + \eta_x \cdot w_y + \zeta_x \cdot w_z \right) -$$

$$\frac{u_0}{f} \left(D_x \cdot w_x + D_x \cdot w_y + D_x \cdot w_z \right) + \frac{u}{f} \left(D_x \cdot w_x + D_x \cdot w_y + D_x \cdot w_z \right) = -\frac{t_\xi}{\delta_u}$$

(30)

By means of equation (29), it is possible to write:

$$D_x \cdot w_x + D_x \cdot w_y + D_x \cdot w_z =$$

$$w_x \left(\xi_x \cdot n_{r,\xi} + \xi_y \cdot n_{r,\eta} + \xi_z \cdot n_{r,\varsigma} \right) + w_y \left(\eta_x \cdot n_{r,\xi} + \eta_y \cdot n_{r,\eta} + \eta_z \cdot n_{r,\varsigma} \right) +$$

$$w_z \left(\zeta_x \cdot n_{r,\xi} + \zeta_y \cdot n_{r,\eta} + \zeta_z \cdot n_{r,\varsigma} \right) +$$

$$\left(t_\xi \cdot n_{r,\xi} + t_\eta \cdot n_{r,\eta} + t_\varsigma \cdot n_{r,\varsigma} \right)$$

If we define the elements:

$$N_\xi = \left(\xi_x \cdot n_{r,\xi} + \xi_y \cdot n_{r,\eta} + \xi_z \cdot n_{r,\varsigma} \right);$$
$$N_\eta = \left(\eta_x \cdot n_{r,\xi} + \eta_y \cdot n_{r,\eta} + \eta_z \cdot n_{r,\varsigma} \right);$$
$$N_\zeta = \left(\zeta_x \cdot n_{r,\xi} + \zeta_y \cdot n_{r,\eta} + \zeta_z \cdot n_{r,\varsigma} \right);$$
$$k = \left(t_\xi \cdot n_{r,\xi} + t_\eta \cdot n_{r,\eta} + t_\varsigma \cdot n_{r,\varsigma} \right),$$

equation (30) becomes:

$$\frac{1}{\delta_u} \left(\xi_x \cdot w_x + \eta_x \cdot w_y + \zeta_x \cdot w_z \right) + \frac{u - u_0}{f} \cdot \left(N_\xi \cdot w_x + N_\eta \cdot w_y + N_\zeta \cdot w_z + k \right) = -\frac{t_\xi}{\delta_u} \Rightarrow$$

$$\left(\frac{\xi_x}{\delta_u} - \frac{(u - u_0) \cdot N_\xi}{f} \right) \cdot w_x + \left(\frac{\eta_x}{\delta_u} - \frac{(u - u_0) \cdot N_\eta}{f} \right) \cdot w_y + \left(\frac{\zeta_x}{\delta_u} - \frac{(u - u_0) \cdot N_\xi}{f} \right) \cdot w_z + \frac{u - u_0}{f} \cdot k =$$

$$-\frac{t_\xi}{\delta_u}$$

(31)

An analogous relation can be written for second equation of (20):

$$\left(\frac{\xi_y}{\delta_v}-\frac{(v-v_0)\cdot N_\xi}{f}\right)\cdot w_x+\left(\frac{\eta_y}{\delta_v}-\frac{(v-v_0)\cdot N_\eta}{f}\right)\cdot w_y+\left(\frac{\zeta_y}{\delta_v}-\frac{(v-v_0)\cdot N_\zeta}{f}\right)\cdot w_z+\frac{v-v_0}{f}\cdot k=-\frac{t_\eta}{\delta_v} \tag{32}$$

By arranging equation (31) and (32), it is possible to redefine the stereoscopic problem, expressed by equation (28):

$$P(u,u,u',v')\cdot w=S \tag{33}$$

In equation (33) P is a matrix 4x3, whose elements depend only by (u,v) and (u',v'), and B is a vector 4x1, whose elements contain parameters of cameras configuration.

The expression of matrix P is:

$$P=\begin{bmatrix} \dfrac{\xi_x}{\delta_u}-\dfrac{(u-u_0)\cdot N_\xi}{f} & \dfrac{\eta_x}{\delta_u}-\dfrac{(u-u_0)\cdot N_\eta}{f} & \dfrac{\zeta_x}{\delta_u}-\dfrac{(u-u_0)\cdot N_\zeta}{f} \\[2mm] \dfrac{\xi_y}{\delta_v}-\dfrac{(v-v_0)\cdot N_\xi}{f} & \dfrac{\eta_y}{\delta_v}-\dfrac{(v-v_0)\cdot N_\eta}{f} & \dfrac{\zeta_y}{\delta_v}-\dfrac{(v-v_0)\cdot N_\xi}{f} \\[2mm] \dfrac{\xi'_x}{\delta_{u'}}-\dfrac{(u'-u'_0)\cdot N_{\xi'}}{f'} & \dfrac{\eta'_x}{\delta_{u'}}-\dfrac{(u'-u'_0)\cdot N_{\eta'}}{f'} & \dfrac{\zeta'_x}{\delta_{u'}}-\dfrac{(u'-u'_0)\cdot N_{\xi'}}{f'} \\[2mm] \dfrac{\xi'_y}{\delta_{v'}}-\dfrac{(v'-v'_0)\cdot N_{\xi'}}{f'} & \dfrac{\eta'_y}{\delta_{v'}}-\dfrac{(v'-v'_0)\cdot N_{\eta'}}{f'} & \dfrac{\zeta'_y}{\delta_{v'}}-\dfrac{(v'-v'_0)\cdot N_{\xi'}}{f'} \end{bmatrix}$$

The expression of vector S is:

$$S=\left\{\begin{array}{c} -\dfrac{t_\xi}{\delta_u}-\dfrac{u-u_0}{f}\cdot k \\[2mm] -\dfrac{t_\eta}{\delta_v}-\dfrac{v-v_0}{f}\cdot k \\[2mm] -\dfrac{t_{\xi'}}{\delta_{u'}}-\dfrac{u'-u'_0}{f'}\cdot k' \\[2mm] -\dfrac{t_{\eta'}}{\delta_{v'}}-\dfrac{v'-v'_0}{f'}\cdot k' \end{array}\right\}$$

By equation (33) it is possibile to invert the problem that is described by eqs.(20) and to recognise the position of a generic point starting to its camera image, by means of pseudoinverse matrix P+ of matrix P.

$$P\cdot w=S \Rightarrow P^T\cdot P\cdot w=P^T\cdot S \Rightarrow w=(P^T\cdot P)^{-1}\cdot P^T\cdot S \Rightarrow w=P^+\cdot S \tag{34}$$

By means of equation (34), it is possible to solve the stereoscopic problem in all configurations in which is verified the condition:

4. The camera calibration

In order to determine the coordinate transformation between the camera reference system and robot reference system, it is necessary to know the parameters that regulate such transformation. The direct measure of these parameters is a difficult operation; it is better to identify them through a procedure that utilize the camera itself.

Camera calibration in the context of three-dimensional machine vision is the process of determining the internal camera geometric and optical characteristics (intrinsic parameters) and/or the 3-D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters). In many cases, the overall performance of the machine vision system strongly depends on the accuracy of the camera calibration.

In order to calibrate the tele-cameras a toolbox, developed by Christopher Mei, INRIA Sophia-Antipolis, was used. By means of this toolbox it is possible to find the intrinsic and extrinsic parameters of two cameras that are necessary to solve the stereoscopic problem. In order to carry out the calibration of a camera, it is necessary to acquire any number of images of observed space in which a checkerboard pattern is placed with different positions and orientations, fig 7.
In each acquired image, after clicking on the four extreme corners of a checkerboard pattern rectangular area, a corner extraction engine includes an automatic mechanism for counting the number of squares in the grid. This points are used like calibration points, fig. 8.
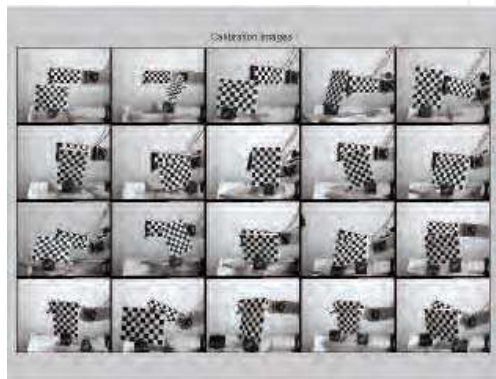
Fig. 7. Calibration images.

The dimensions **dX**, **dY** of each of squares are always kept to their original values in millimeters, and represent the parameters that put in relation the pixel dimensions with observed space dimensions (mm).
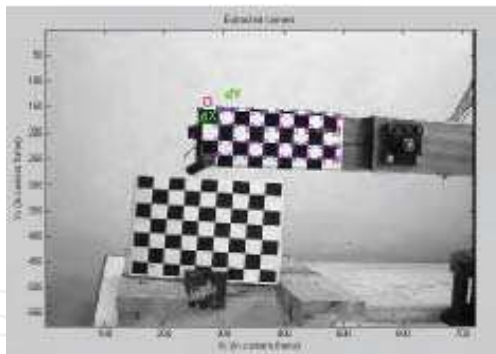

Fig. 8. Calibration image After corner extraction, calibration is done in two steps: first initialization, and then nonlinear optimization.

The initialization step computes a closed-form solution for the calibration parameters based not including any lens distortion.
The non-linear optimization step minimizes the total reprojection error (in the least squares sense) over all the calibration parameters (9 DOF for intrinsic: focal (2), principal point (2), distortion coefficients (5), and 6*n DOF extrinsic, with n = images number).
The calibration procedure allows to find the 3-D position of the grids with respect to the camera, like shown in fig. 9.
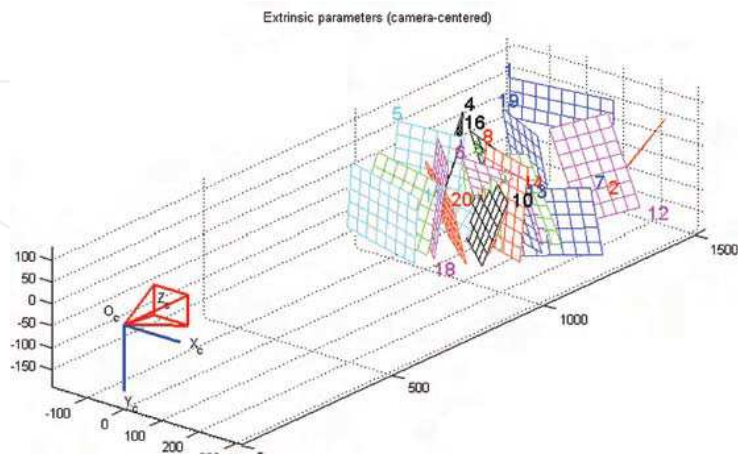
Fig. 9. Position of the grids for the calibration procedure.

With two camera calibration, it is possible to carry out a stereo optimization, by means of a toolbox option, that allows to do a stereo calibration for stereoscopic problem.

The global stereo optimization is performed over a minimal set of unknown parameters, in particular, only one pose unknown (6 DOF) is considered for the location of the calibration grid for each stereo pair. This insures global rigidity of the structure going from left view to right view. In this way the uncertainties on the intrinsic parameters (especially that of the focal values) for both cameras it becomes smaller.

After this operation, the spatial configuration of the two cameras and the calibration planes may be displayed in a form of a 3D plot, like shown in fig. 10.
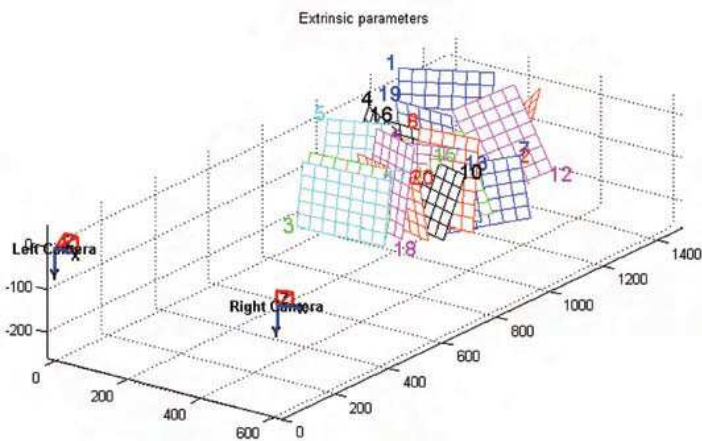

Fig.10. Calibration planes.

5. Robot cinematic calibration

Among the characteristics that define the performances of a robot the most important can be considered the repeatability and the accuracy. Generally, both these characteristics depend on factors like backlashes, load variability, positioning and zero putting errors, limits of the

transducers, dimensional errors, and so on. The last sources of error essentially depend on the correct evaluation of the Denavit and Hartemberg parameters. Hence, some of the sources of error can be limited by means of the cinematic calibration.

Basically, by the cinematic calibration it is assumed that if the error in the positioning of the robot's end effector is evaluated in some points of the working space, by means of these errors evaluation it is possible to predict the error in any other position thus offset it.

In few words, the main aim of the technique showed in this paper is to obtain precise evaluations of those Denavit-Hartenberg parameters that represent, for each of the links, the length, the torsion and the offset.

5.1 The calibration technique

This calibration technique essentially consists in the following steps:

I.   The end-effector is located in an even position in the work space;

II. A vision system acquires and records the robot's image and gives the coordinates of an assigned point of the end-effector, expressed in pixels in the image plane.

III. By means of a suitable camera model, it is possible to find a relation between these coordinates expressed in pixels, and the coordinates of the assigned point of the end-effector in the world (Cartesian) frame.

IV. By means of the servomotor position transducers, the values of the joint position parameters are recorded for that end-effector position in the work space. In this way, for each of the camera images, the following arrays are obtained:

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}, \quad \begin{pmatrix} \theta_{1,i} \\ \theta_{2,i} \\ \theta_{3,i} \end{pmatrix} \tag{35}$$

where: $i = 1,\ldots,N$, and N is the number of acquired camera images (frames).

If the coordinates in the working space and the joint parameters are known, it's possible to write the direct kinematics equations in which the unknown are those Denavit-Hartenberg parameters that differ from the joint parameters; thus these Denavit-Hartenberg parameters represent the unknown of the kinematic calibration problem.

The expression of these equations is obtained starting from the transform matrix (homogeneous coordinates) that allows to transform the coordinates in the frame i to the coordinates in the frame i-1:

$$^{i-1}A_i = \begin{bmatrix} C\vartheta_i & -C\alpha_i \cdot S\vartheta_i & S\alpha_i \cdot S\vartheta_i & a_i \cdot C\vartheta_i \\ S\vartheta_i & C\alpha_i \cdot C\vartheta_i & -S\alpha_i \cdot C\vartheta_i & a_i \cdot S\vartheta_i \\ 0 & S\alpha_i & C\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{36}$$

By means of such matrixes it is possible to obtain the transform matrix that allows to obtain the coordinates in the frame 0 (the fixed one) from those in frame n (the one of the last link) :

$$^0T_n = {}^0A_1 \cdot {}^1A_2 \cdot \ldots \cdot {}^{n-1}A_n .$$

As for an example, if we consider a generic 3 axes revolute (anthropomorphic) robot arm, we'll obtain an equation that contains 9 constant kinematic parameters and 3 variable parameters ($\theta_1, \theta_2, \theta_3$).

So, the vector:

$$\pi_{DH} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ d_1 \\ d_2 \\ d_3 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \qquad (37)$$

represents the unknown of the kinematic calibration problem.
Said:

$$\Theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \qquad (38)$$

the direct kinematics equation for this manipulator can be written as :

$$w = t_4(\pi_{DH}, \Theta) \qquad (39)$$

where w is the position vector in the first frame and is the fourth row of the Denavit-Hartenberg transform matrix. In equation (39) it clearly appears that the position depends on the joint parameters and on the others Denavit-Hartenberg parameters. Equation (39) can be also seen as a system of 3 equations (in Cartesian coordinates) with 9 unknowns: the elements of vector π_{DH} .

Obviously, it's impossible to solve this system of equations, but it's possible to use more camera images taken for different end-effector positions:

$$\begin{cases} t_4(\pi_{DH}, \Theta^1) = w_1 \\ t_4(\pi_{DH}, \Theta^2) = w_2 \\ \dots\dots\dots\dots\dots \\ t_4(\pi_{DH}, \Theta^N) = w_N \end{cases} \qquad (40)$$

with $N \geq 9$.

As, for each of the camera images the unknown Denavit-Hartemberg parameters are the same, equations (40) represent a system of N non linear equations in 9 unknowns. This system can be numerically solved by means of a minimum square technique.

It's known at a minimum square problem can be formulated as follows:
given the equation (39), find the solutions that minimize the expression:

$$\int_{D_\Theta} |t_4(\pi_{DH}, \Theta) - w|^2 \cdot d\Theta \qquad (41)$$

This method can be simplified by substituting the integrals with summations, thus it must be computed the vector that minimize the expression:

$$\sum_{i=1}^{N} |t_4(\pi_{DH}, \Theta^i) - w_i|^2 \qquad (42)$$

If we formulate the problem in this way, the higher is the number of images that have been taken (hence the more are the known parameters), the more accurate will be the solution, so it's necessary to take a number of pictures.

6. Trajectories recording

The trajectory recording, that is essential to study robot arm dynamical behaviour has been obtained by means of two digital television camera linked to a PC.

The rig, that has been developed, is based on a couple of telecameras; it allows us to obtain the velocity vector of each point of the manipulator. By means of this rig it is possible:

- to control the motion giving the instantaneous joint positions and velocities;
- to measure the motions between link and servomotor in presence of non-rigid transmissions;
- to identify the robot arm dynamical parameters.

An example of these video application for robot arm is the video acquisition of a robot arm trajectories in the work space by means of the techniques above reported.

In the figures 11 and 12 are reported a couple of frames, respectively, from the right telecamera and the left one. In fig. 13 is reported the 3-D trajectory, obtained from the frames before mentioned; in this last figure, for comparison, the trajectory obtained from the encoders signals is also reported.

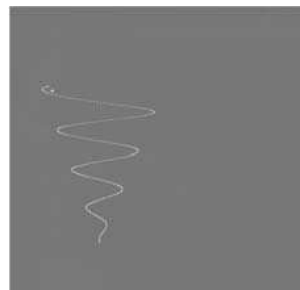Fig. 11. Trajectories in image space: camera position 1.

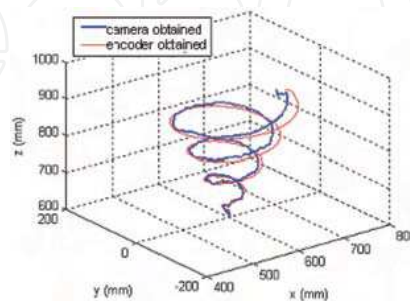Fig. 12. Trajectories in image space: camera position 2.

Fig. 13. comparison between trajectory recordings.

7. References

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

V. Niola, C.Rossi: A method for the trajectories recording of a robot arm: early results -*Proc. 9th WSEAS International Conference on Computers*, Athens 11 – 16 July 2005.

V.Niola, C. Rossi: Video acquisition of a robot arm trajectories in the work space - *WSEAS Transactions on Computers*, Iusse 7, Vol. 4, july 2005, pagg. 830 – 836.

G.Nasti, V.Niola, S.Savino: Experimental Study on Vision Motion Measures in Robotics -*9th WSEAS International Conference on Computers*, Vouliagmeni – Athens, Greece, July 11 – 16, 2005.

V. Niola, C. Rossi, S. Savino: Modelling and Calibration of a Camera for Trajectories Recording -*Proc. 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation* – Madrid, February 15-17, 2006

V. Niola, C. Rossi, S. Savino: A Robot Kinematic Calibration Technique -*Proc. 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation* – Madrid, February 15-17, 2006

A. Fusiello: Visione Computazionale: appunti delle lezioni – Informatic Department, University of Verona, 3 March 2005.

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.

R. Sharma –"Active vision for visual servoing: a review - *IEEE Workshop on Visual Servoing: Achivement"* - Application and Open Problems, Maggio 1994.

R. Sharma, S, Hutchinson – "Optimizing hand/eye configuration for visual-servo system" - *IEEE International Confeence on Robotics and Automation* , pp. 172-177, 1995.

C. Mei: Camera Calibration Toolbox for Matlab -http://www.vision.caltech.edu/ bouguetj/calib_doc.

V. Niola, C. Rossi, S. Savino – "Modelling and calibration of a camera for robot trajectories recording" -*Proc. 5th WSEAS Int. Conf. on "Signal Processing, Robotics and Automation"* – Madrid, February 15-17, 2006.

A. Fusiello – "Visione Computazionale: appunti delle lezioni" – Informatic Department, University of Verona, 3 March 2005.

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

R. Brancati, C. Rossi, S. Savino, A Method for Trajectory Planning in the Joint Space, *Proc. of 14th International Workshop on Robotics in Alpe-Adria-Danube Region,* Bucharest, Romania, May 26-28, 2005, pp.81-85.

R. Sharma, S. Hutchinson, On the observability of robot motion under active camera control, *Proc. IEEE International Conference on Robotics and Automation,* May 1999, pp. 162-167.

J. T. Feddema, C. S. George Lee, O. R. Mitchell, Weighted selection of image features for resolved rate visual feedback control, *IEEE Trans. Robot. Automat.*, Vol. 7, 1991, pp. 31-47.

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.
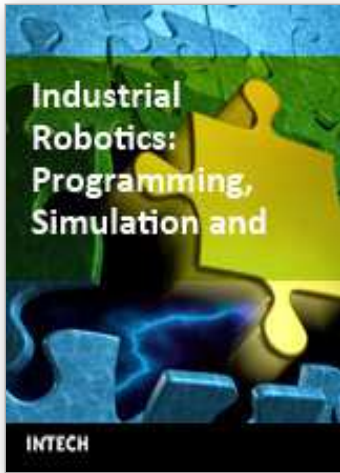
Industrial Robotics: Programming, Simulation and Applications
Edited by Low Kin Huat

ISBN 3-86611-286-6
Hard cover, 702 pages
Publisher Pro Literatur Verlag, Germany / ARS, Austria
Published online 01, December, 2006
Published in print edition December, 2006

This book covers a wide range of topics relating to advanced industrial robotics, sensors and automation technologies. Although being highly technical and complex in nature, the papers presented in this book represent some of the latest cutting edge technologies and advancements in industrial robotics technology. This book covers topics such as networking, properties of manipulators, forward and inverse robot arm kinematics, motion path-planning, machine vision and many other practical topics too numerous to list here. The authors and editor of this book wish to inspire people, especially young ones, to get involved with robotic and mechatronic engineering technology and to develop new and exciting practical applications, perhaps using the ideas and concepts presented herein.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Vincenzo Niola and Cesare Rossi (2006). Video Applications for Robot Arms, Industrial Robotics: Programming, Simulation and Applications, Low Kin Huat (Ed.), ISBN: 3-86611-286-6, InTech, Available from: http://www.intechopen.com/books/industrial_robotics_programming_simulation_and_applications/video_applications_for_robot_arms

INTECH
open science | open minds