Proceedings of the 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation, Madrid, Spain, February 15-17, 2006 (pp87-91)

Perspective Transform and Vision System for Robotic Applications

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: - A vision algorithm is proposed by means of which it is 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.

Key- words: Vision system, Robot trajectories, Perspective transform, perspective image, robot matrix transformation.

1 Introduction

By means of perspective transformation it is possible to associate a point in the geometric space to a point in a plane. In homogeneous coordinates the perspective transformation matrix has non-zero elements in the fourth row. An espression of perspective transformation is proposed with the scope to introduce the perspective concepts for the application in robotic field.

2- The vision algorithm

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 perspective transformation matrix

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 **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.

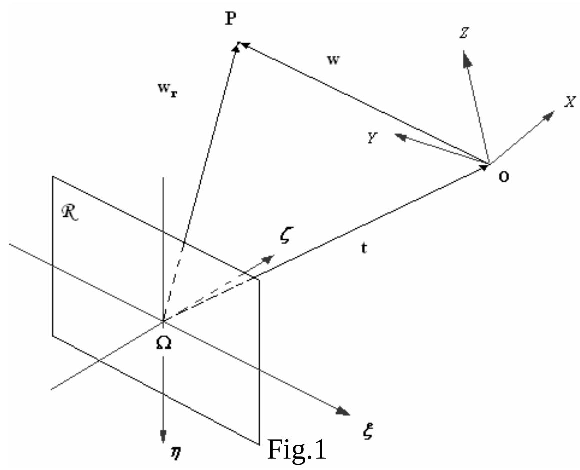


Fig.1

The vectors above are joined by the equation:

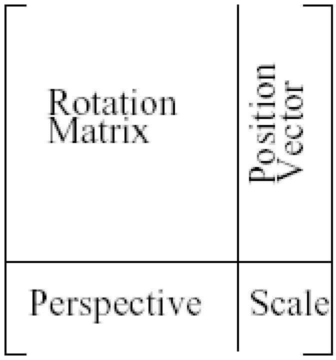
$$\begin{Bmatrix} w_{r,x} \\ w_{r,y} \\ w_{r,z} \\ sf \end{Bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_{\xi} \\ R_{21} & R_{22} & R_{23} & t_{\eta} \\ R_{31} & R_{32} & R_{33} & t_{\zeta} \\ 0 & 0 & 0 & sf \end{bmatrix} = \begin{Bmatrix} w_{x} \\ w_{y} \\ w_{z} \\ sf \end{Bmatrix} \qquad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

where sf is the scale factor; more concisely eq.1 can be written as follows:

$$\tilde{w}_r = T \cdot \tilde{w} \qquad (2)$$

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:



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 $\mathbf{w_r}$ on the plane ξ,η.

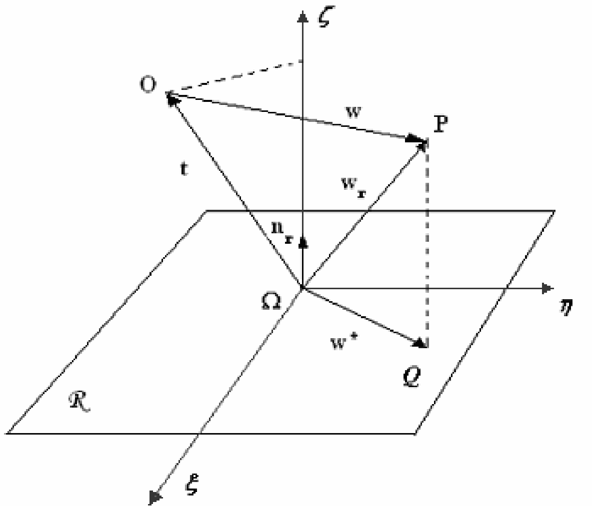


Fig.2

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

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

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:

$$\tilde{w}^* = \begin{pmatrix} w_{r,\xi} \\ w_{r,\eta} \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \hat{\xi}^T w + t_\xi \\ \hat{\eta}^T w + t_\eta \\ 0 \\ 1 \end{pmatrix} \qquad (4)$$

In the same figure, $\mathbf{n_r}$ 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:

$$d = n_r^{\,T} w_r \qquad (5)$$

Let's indicate with $\mathbf{w}_{\{\Omega,\xi,\eta,\zeta\}}$ the vector **w** in the frame {Ω,ξ,η,ζ}:

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

Because $\hat{\xi}\ \ \hat{\eta}\ \ \hat{\zeta}$ are the versor of the frame {Ω,ξ,η,ζ} 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 {Ω,ξ,η,ζ}:

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

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

$$\tilde{w}_r = \tilde{w}_{\{\Omega,\xi,\eta,\zeta\}} + \tilde{t} = \begin{pmatrix} w_\xi + t_\xi \\ w_\eta + t_\eta \\ w_\zeta + t_\zeta \\ 1 \end{pmatrix} =$$
$$= \begin{pmatrix} \xi_x w_x + \xi_y w_y + \xi_z w_z + t_\xi \\ \eta_x w_x + \eta_y w_y + \eta_z w_z + t_\eta \\ \zeta_x w_x + \zeta_y w_y + \zeta_z w_z + t_\zeta \\ 1 \end{pmatrix} \qquad (5')$$

an expression of d is:

Proceedings of the 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation, Madrid, Spain, February 15-17, 2006 (pp87-91)

$$d = n_r^{\ T} w_r = \begin{pmatrix} n_{r,\xi} \\ n_{r,\eta} \\ n_{r,\zeta} \end{pmatrix}^T \cdot \begin{pmatrix} w_\xi + t_\xi \\ w_\eta + t_\eta \\ w_\zeta + t_\zeta \end{pmatrix} =$$

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

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 possibile to write:

$$d = n_r^{\ T} w_r = \begin{pmatrix} D_x \\ D_y \\ D_z \\ 0 \end{pmatrix}^T \cdot \begin{pmatrix} w_x \\ w_y \\ w_z \\ 1 \end{pmatrix} = D^T \cdot w \qquad (5")$$

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:

$$\tilde{w}*_p = \begin{pmatrix} \hat{\xi}^T w + t_\xi \\ \hat{\eta}^T w + t_\eta \\ 0 \\ n_r^{\ T} w_r \end{pmatrix} \qquad (6)$$

The perspective matrix [Tp] can be obtained:

$$\tilde{w}*_p = T_p \cdot \tilde{w} \qquad \Rightarrow$$

$$T_p = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ D_x & D_y & D_z & 0 \end{bmatrix} \qquad (7)$$

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

3- The perspective concept

From eq. 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 Δw, the point moves from the position P to the position P', that are given by the vectors:

$$w = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} \quad \text{e} \quad w' = \begin{pmatrix} w'_x \\ w_y \\ w_z \end{pmatrix} \qquad (8)$$

If the perspective transforms are applied we have :

$p = T_p \cdot w$ and $p' = T_p \cdot w'$

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

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

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

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

that is to say:

$$\Delta p = \begin{pmatrix} \dfrac{\xi_x \cdot [w_x'(D^T w) - w_x (D'^T\ w')]}{(D^T w)(D'^T\ w')} \\ \dfrac{\eta_x \cdot [w_x'(D^T w) - w_x (D'^T\ w')]}{(D^T w)(D'^T\ w')} \\ 0 \end{pmatrix}$$

(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 :

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

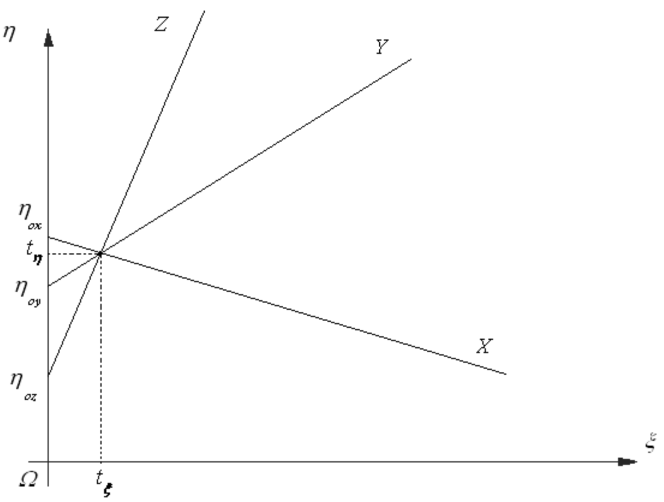


Fig.3

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:

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

z axis : $\eta = (\eta_z / \xi_z) \cdot \xi + \dfrac{\xi_z t_\eta - \eta_z t_\xi}{\xi_z}$ (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 eq. (14) to the corresponding one in the image plane; in fact by means of equations (11),(12) e (13) it is possible to write:

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

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.

4- Robotic Application

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:

$$A_1^{i-1} = \begin{bmatrix} C\vartheta_i & -C\alpha_i \cdot S\vartheta_i & S\alpha_i \cdot S\vartheta_i & a_i \cdot C\vartheta_i \\ S\vartheta_i & C\alpha_i \cdot C\vartheta_i & -S\alpha_i \cdot C\vartheta_i & a_i \cdot S\vartheta_i \\ 0 & S\alpha_i & C\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

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

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

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 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:

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

where $\{P\}_p$ is the perspective image of generic point P and $\left[T_p\right]_n^0$ 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.

5- Conclusion

An espression of perspective transformation is proposed with the scope to introduce the perspective concepts for the application in robotic field. A 4x4 transformation matrix is shown and the forth row represents the perspective transformation that is the same of Denavit and Hartemberg matrix in a generic robot structure.

The expression of perspective transformation matrix from robot end-effector frame to an image plane, will be used in robotic vision system, to “recognize” robot behavior by means of an adapted camera system.

Aknowledgments

The authors thanks Mr. Vincenzo Ciccarelli, Mr. Giovanni D’Orsi and Mr. Angelo Proni that gave an useful contribution on the researches on these topics while working for their thesis for the master.

References

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

[2] 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

[3] 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.

[4] 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.

[5] 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 Confeence on Robotics and Automation , pp. 172-177, 1995.

[10] S. SAVINO – Progettazione, realizzazione e messa a punto di un robot revoluto. – PhD Thesis in Mechanical Systems Engineering, VII, A.A. 2003-04, D.I.M.E. Univ. Of Naples.