

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

4,600

Open access books available

120,000

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



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



Video System in Robotic Applications

Vincenzo Niola, Cesare Rossi, Sergio Savino and Salvatore Strano
University of Naples "Federico II"
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 images for automation purposes is an interesting topic.

Industrial applications 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 these techniques 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.

The use of a vision system in a robot application, it concurs to increase the robots ability to interact with their work space, to make more efficient their management.

In this chapter some "Artificial Vision" applications to robotics are described:

- robot cinematic calibration;
- trajectories recording;
- path planning by means of vision system;
- solid reconstruction with a video system on a robot arm.

2. Vision usefulness

The man perceives the characteristics of the external world by means of sense organs. They allow that a sure flow of information, regarding for example the shape, the color, the temperature, the smell of an object, reaches the brain; in this way each man possesses a complete description of that is around him. More in a generalized manner, the man can itself be seen as a system that, for survival reasons, must interact with the external world,

and, to be able to make it, he has need of a sensory apparatus able to supply him continuously information. This affirmation can be extended also to not biological systems, like, for example, the automatic machines or the robots. These equipment, carrying out a determined task, interact with the external world and they must be fortified with devices that are able to perceive the world characteristics, these devices are called sensors. A robot or one whichever automatic machine, that is equipped with sensors, is able to perceive the "stimula" of the external world, in which it works.

What is the difference between sensors and organs of sense, at the operating level?

When we perceive a any sound, for example the voice of a known person, we are able to distinguish the stamp, to establish if it is acute or serious, to feel the intensity and volume. If instead, the same voice is acquired through a microphone, converting its signal in digital and processing it by means of an electronic calculator, the information that we can deduce, increase and they make more detailed: we will be able, for example, to determine the main frequency components, to measure the amplitude in decibel, to visualize the wave shape. In other words, sensors, beyond to having the representative function of the truth, concur also to extrapolate information at quantitative level and they allow us to lead a technical analysis on the acquired data.

Main aim of this chapter is to show how it is possible to equip robot of sight

The main problem of the reliable and precise robot realization, has been to implement the hardware e software structures, that constitute a sturdy and efficient control system.

How does motion control work in the man? The human body has a much elevating number of degrees of freedom and this renders very arduous the nervous system task. For this reason a highly centralized control structure is necessary. It is possible to describe the job of such structure through a simple example: let's image a man, that wants to take an object that is disposed on a table distant some meters. The man observes the table and the object position, while the brain elaborates the trajectory and the nervous impulses to transmit to muscles, characterizing reference points that are acquired from image observed by eyes. Subsequently the man begins to move and after some step, he reaches the table and takes the object. From this example, it can be asserted that, excluding the memory contribution and an elevated development of the other senses, it is not possible to carry out a task without to see.

Therefore the sense of the sight it has a twofold function in the human body motion:

1. to characterize the targets in the space.
2. to control the position and the guidelines of the several parts of the body that move.

In the robots, the motion control is, usually implemented only with joints position transducers. It is clear that, if joints translation and rotations are known, its spatial configuration is known completely; therefore, the second function that has been attributed to human sight is realized. A blind control is less suitable to catch up a target in the work space. In fact, to guide the robot end effector to a point, it is necessary to know, with precision, its cartesian coordinates and "translate" them in the joints space by means of inverse cinematic. For this reason, it is useful to increase robots sensory abilities, equipping them "off sight", by means of vision systems with opportune sensors. In this chapter it will be described in how it is possible to characterize the targets in the work space and to determine the values of the Denavit-Hartenberg cinematic parameters, by means of opportune techniques, and with two television cameras, so as to make simpler, more accurate and efficient both the robot management and the motion planning.

3. Vision process

About the term “vision” applications to industrial robots, the meaning of this word must be enriched and cleared with technical slight knowledge.

In literature, use of the vision like instrument for technical applications, is called “machine vision” or “computer vision”.

It is important to explain, in the first place, which is the aim of the computer vision: to recognize the characteristics of objects that are present in the acquired images of work space and to associate them their real meant.

The vision process can be divided in following operations:

- Perception
- Pre-elaboration
- Segmentation
- Description
- Recognition
- Interpretation

Perception is the process that supplies the visual image. With this operation, we mean the mechanism of photogram formation by means of a vision system and a support, like a computer.

Pre-elaboration is the whole of noise reduction techniques and images improvement techniques.

Segmentation is the process by means of which the image is subdivided in characteristics of interest.

Description carries out the calculation of the characteristics that segmentation has evidenced, it represents the phase in which it is possible to quantify that only qualitatively has been characterized: lengths, areas, volumes, ecc.

Recognition consists in assembling all the characteristics that belong to the object, in order to characterize the object. By means of the last phase, called interpretation, it is established the effective correspondence between a characterized shape and the object that is present in the real scene.

To say that a robot “sees”, does not mean simply that it has a reality representation, but that it is able to recognize quantitatively the surrounding space, that is to recognize distances, angles, areas and volumes of the objects that are in the observed scene.

4. 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 Hartenberg 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.

4.1 The matrix for the perspective transformation [1,5]

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 w_r and the image plane is indicated with R ; this last, for the sake of simplicity is supposed to be coincident with the plane ξ,η .

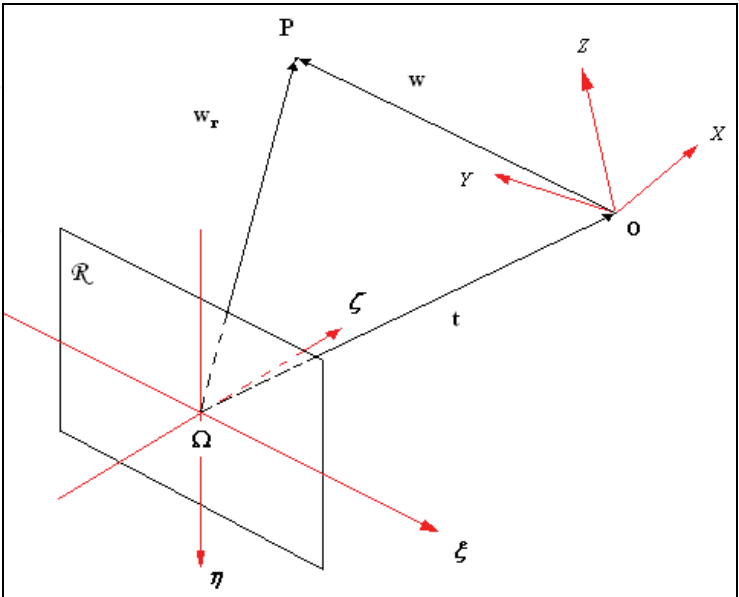


Figure 1. Frames for the perspective transformation

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:

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

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 prospective transform three values, generally different by zero, will be determined.

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

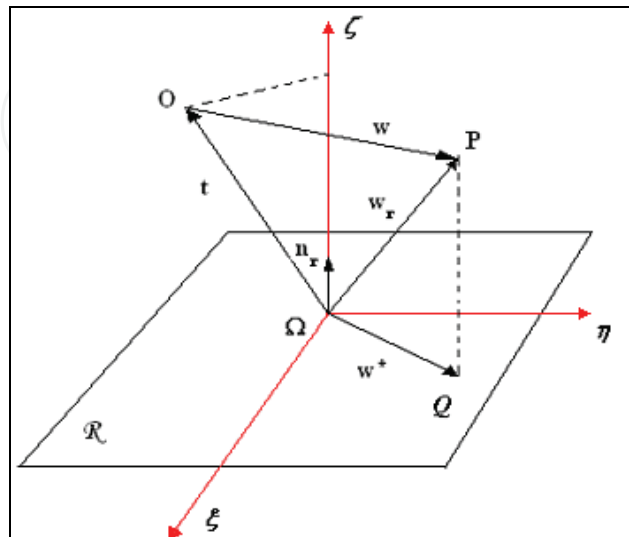


Figure 2. Vectors for the perspective transformation

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:

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

where $\hat{\xi} \ \hat{\eta} \ \hat{\zeta}$ are the versor of the frame $\{\Omega, \xi, \eta, \zeta\}$ 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:

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

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:

$$d = \mathbf{n}_r^T \mathbf{w}_r \quad (5)$$

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

$$\tilde{\mathbf{w}}_{\{\Omega, \xi, \eta, \zeta\}} = \begin{pmatrix} w_\xi \\ w_\eta \\ w_\zeta \\ 1 \end{pmatrix} \quad (6)$$

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 $\{\Omega, \xi, \eta, \zeta\}$:

$$\begin{aligned} w_\xi &= \hat{\xi}^T \cdot \mathbf{w} = \xi_x w_x + \xi_y w_y + \xi_z w_z; \\ w_\eta &= \hat{\eta}^T \cdot \mathbf{w} = \eta_x w_x + \eta_y w_y + \eta_z w_z; \\ w_\zeta &= \hat{\zeta}^T \cdot \mathbf{w} = \zeta_x w_x + \zeta_y w_y + \zeta_z w_z; \end{aligned} \quad (7)$$

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

$$\tilde{\mathbf{w}}_r = \tilde{\mathbf{w}}_{\{\Omega, \xi, \eta, \zeta\}} + \tilde{\mathbf{t}} = \begin{pmatrix} w_\xi + t_\xi \\ w_\eta + t_\eta \\ w_\zeta + t_\zeta \\ 1 \end{pmatrix} \quad (8)$$

Let's introduce the expressions:

$$\begin{aligned} D_x &= \frac{(\xi_x w_x + \xi_y w_y + \xi_z w_z + t_\xi) \cdot \mathbf{n}_{r, \xi}}{w_x}; \\ D_y &= \frac{(\eta_x w_x + \eta_y w_y + \eta_z w_z + t_\eta) \cdot \mathbf{n}_{r, \eta}}{w_y}; \\ D_z &= \frac{(\zeta_x w_x + \zeta_y w_y + \zeta_z w_z + t_\zeta) \cdot \mathbf{n}_{r, \zeta}}{w_z}; \end{aligned} \quad (9)$$

it is possible to write:

$$d = n_r^T w_r = \begin{pmatrix} D_x \\ D_y \\ D_z \\ 0 \end{pmatrix}^T \cdot \begin{pmatrix} w_x \\ w_y \\ w_z \\ 1 \end{pmatrix} = D^T \cdot w \quad (10)$$

In the equation (10) the vector D is:

$$D = \begin{pmatrix} D_x \\ D_y \\ D_z \\ 0 \end{pmatrix} \quad (11)$$

As vector w^* is given by:

$$\tilde{w}^*_p = \begin{pmatrix} \hat{\xi}^T w + t_\xi \\ \hat{\eta}^T w + t_\eta \\ 0 \\ n_r^T w_r \end{pmatrix} \quad (12)$$

The perspective matrix $[T_p]$ can be obtained:

$$\tilde{w}^*_p = T_p \cdot \tilde{w} \Rightarrow T_p = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ D_x & D_y & D_z & 0 \end{bmatrix} \quad (13)$$

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.

4.2 The perspective concept

From equation (13) 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 what 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:

$$w = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} \quad \text{and} \quad w' = \begin{pmatrix} w'_x \\ w'_y \\ w'_z \end{pmatrix} \quad (14)$$

If the perspective transforms are applied we have :

$$p = T_p \cdot w \quad \text{and} \quad p' = T_p \cdot w' \quad (15)$$

the displacement in the image plane is given by:

$$\Delta p = p' - p \quad (16)$$

that is to say:

$$\Delta p = \begin{pmatrix} \frac{\xi_x \cdot [w'_x (D^T w) - w_x (D'^T w')]}{(D^T w)(D'^T w')} \\ \frac{\eta_x \cdot [w'_x (D^T w) - w_x (D'^T w')]}{(D^T w)(D'^T w')} \\ 0 \end{pmatrix} \quad (17)$$

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 . So the x axis equation in the image plane is:

$$\eta = (\eta_x / \xi_x) \cdot \xi + \frac{\xi_x t_\eta - \eta_x t_\xi}{\xi_x} \quad (18)$$

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

$$y \text{ axis : } \eta = (\eta_y / \xi_y) \cdot \xi + \frac{\xi_y t_\eta - \eta_y t_\xi}{\xi_y} \quad (19)$$

$$z \text{ axis : } \eta = (\eta_z / \xi_z) \cdot \xi + \frac{\xi_z t_\eta - \eta_z t_\xi}{\xi_z} \quad (20)$$

By means of equations (18), (19) and (20) 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.

4.3 Perspective transformation in D-H robotic matrix [5]

For kinematics purposes in robotic applications, it is possible to use the Denavit and Hartenberg 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:

$$A_1^{i-1} = \begin{bmatrix} C\vartheta_1 & -C\alpha_1 \cdot S\vartheta_1 & S\alpha_1 \cdot S\vartheta_1 & a_1 \cdot C\vartheta_1 \\ S\vartheta_1 & C\alpha_1 \cdot C\vartheta_1 & -S\alpha_1 \cdot C\vartheta_1 & a_1 \cdot S\vartheta_1 \\ 0 & S\alpha_1 & C\alpha_1 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (21)$$

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

$$T_n^0 = A_1^0 \cdot A_2^1 \cdot A_3^2 \cdot \dots \cdot A_n^{n-1} \quad (22)$$

With this matrix it is possible to solve the expression:

$$\{P\}_0 = T_n^0 \cdot \{P\}_n \quad (23)$$

where $\{P\}_0$ and $\{P\}_n$ are the vectors that represent a generic point P in frame 0 and frame n .

It is useful to include the perspective concepts in this transformation 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:

$$\{P\}_p = T_p \cdot \{P\}_0 = T_p \cdot T_n^0 \cdot \{P\}_n = \left[T_p \right]_n^0 \cdot \{P\}_n \quad (24)$$

where $\{P\}_p$ is the perspective image of generic point P and T_p is the perspective transformation matrix from end-effector frame to an image plane.

With this representation the fourth row of the Denavit and Hartenberg matrix will contain non-zero elements. A vision system demands an application like this.

5. 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.

5.1 The thin lenses model [2,4,13,14]

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

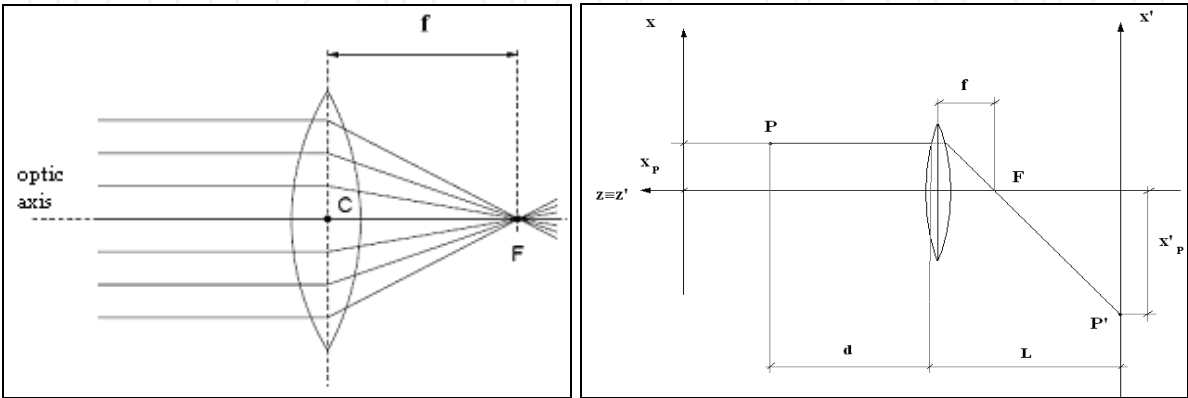


Figure 3. Convergent lens (left), Thin lens (right)

The focal distance f, in air, is given by:

$$f = (n - 1) \cdot \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \tag{25}$$

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.3 the equation for the thin lenses gives:

$$\frac{1}{d} + \frac{1}{L} = \frac{1}{f} \tag{26}$$

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.3).
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:

$$\frac{x_P}{f} = - \frac{x_P}{L - f} \tag{27}$$

with the equation of the thin lenses we can write:

$$x_P' = -\frac{f}{d-f} \cdot x_P \quad (28)$$

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:

$$x_P' \cong -\frac{f}{d} \cdot x_P \quad (29)$$

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

5.2 The model of the camera [4]

As already observed a camera 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 on an image it is possible to define a frame u,v (fig.4) 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 camera. 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 can be revealed by the camera and isn't a normalized length unit. The model of the camera must take into account this aspect also.

In order to obtain the model of the camera the scheme reported in fig.4 can be considered.

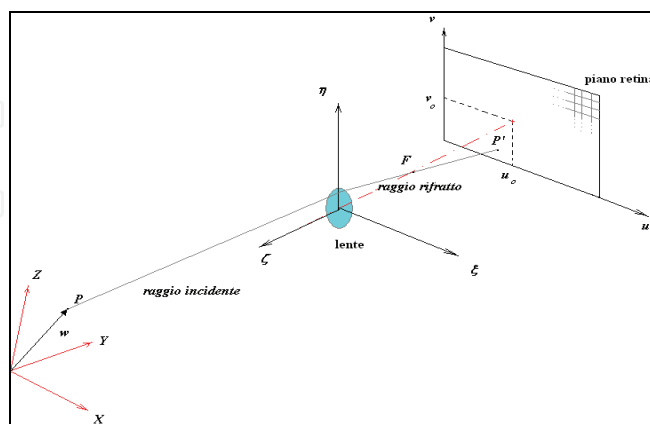


Figure 4. Camera model

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 ξ, η, ζ having the origin in the lens centre and the plane ξ, η coincident with the plane of the lens; hence, the plane ξ, η 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_0 and v_0 are the coordinates of the origin of frame ξ, η, ζ 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_1 vector is obtained:

$$\tilde{w}_1 = T_1 \cdot \tilde{w} \quad (30)$$

Where the matrix T_1 is obtained dividing by $-f$ the last row of the perspective transformation matrix T_p .

$$T_1 = \begin{bmatrix} \xi_x & \xi_y & \xi_z & t_\xi \\ \eta_x & \eta_y & \eta_z & t_\eta \\ 0 & 0 & 0 & 0 \\ -\frac{D_x}{f} & -\frac{D_y}{f} & -\frac{D_z}{f} & 0 \end{bmatrix} \quad (31)$$

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_1 e y_1 as the first two components of the vector w_1 , the coordinates u and v (expressed in pixel) of P' (image of P) are :

$$\begin{cases} u = \frac{x_1}{\delta_u} + u_0 \\ v = \frac{y_1}{\delta_v} + v_0 \end{cases} \quad (32)$$

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

So, by substituting equation (30) in equation (32) it comes:

$$\begin{cases} u = -\frac{f}{D^T w} \left[\left(\frac{1}{\delta_u} \cdot \hat{\xi} - \frac{u_0}{f} \cdot D \right)^T w + \frac{1}{\delta_u} \cdot t_\xi \right] \\ v = -\frac{f}{D^T w} \left[\left(\frac{1}{\delta_v} \cdot \hat{\eta} - \frac{v_0}{f} \cdot D \right)^T w + \frac{1}{\delta_v} \cdot t_\eta \right] \end{cases} \quad (33)$$

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

$\tilde{m} = [m_1 \ m_2 \ -D^T w/f]^T = [u \ v \ -D^T w/f]^T$ of the previous vector can be written :

$$\tilde{m} = M \cdot \tilde{w} \quad (34)$$

Where M is the matrix :

$$M = \begin{bmatrix} \left(\frac{\xi_x}{\delta_u} - \frac{u_o D_x}{f} \right) & \left(\frac{\xi_y}{\delta_u} - \frac{u_o D_y}{f} \right) & \left(\frac{\xi_z}{\delta_u} - \frac{u_o D_z}{f} \right) & t_\xi / \delta_u \\ \left(\frac{\eta_x}{\delta_v} - \frac{v_o D_x}{f} \right) & \left(\frac{\eta_y}{\delta_v} - \frac{v_o D_y}{f} \right) & \left(\frac{\eta_z}{\delta_v} - \frac{v_o D_z}{f} \right) & t_\eta / \delta_v \\ -D_x / f & -D_y / f & -D_z / f & 0 \end{bmatrix} \quad (35)$$

that represents the requested model of the camera.

6. The stereoscopic vision

What 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 equations (33) becomes a system of 2 equation with 3 unknowns, so it has no solutions.

This obstacle can be overcome by means of a vision system with at least two cameras.

In this way, what 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 (34) we have:

$$\begin{cases} \tilde{m} = M \cdot w \\ \tilde{m}' = M' \cdot w \end{cases} \quad (36)$$

The first equation of the system (36), in Cartesian coordinates (non-homogenous), can be written as:

$$\begin{cases} (u \cdot D + f \cdot \mu_1)^T w = \mu_{14} \\ (v \cdot D + f \cdot \mu_2)^T w = \mu_{24} \end{cases} \quad (37)$$

Where:

$$\begin{aligned} \mu_1 &= \left\{ \left(\frac{\xi_x}{\delta_u} - \frac{u_o D_x}{f} \right) \left(\frac{\xi_y}{\delta_u} - \frac{u_o D_y}{f} \right) \left(\frac{\xi_z}{\delta_u} - \frac{u_o D_z}{f} \right) \right\}; \\ \mu_2 &= \left\{ \left(\frac{\eta_x}{\delta_v} - \frac{v_o D_x}{f} \right) \left(\frac{\eta_y}{\delta_v} - \frac{v_o D_y}{f} \right) \left(\frac{\eta_z}{\delta_v} - \frac{v_o D_z}{f} \right) \right\}; \\ \mu_{14} &= t_\xi / \delta_u \\ \mu_{24} &= t_\eta / \delta_v \end{aligned} \quad (38)$$

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

$$\begin{cases} (u' \cdot D' + f' \cdot \mu'_1)^T w = \mu'_{14} \\ (v' \cdot D' + f' \cdot \mu'_2)^T w = \mu'_{24} \end{cases} \quad (39)$$

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

$$\begin{bmatrix} (u \cdot D + f \cdot \mu_1)^T \\ (v \cdot D + f \cdot \mu_2)^T \\ (u' \cdot D' + f' \cdot \mu'_1)^T \\ (v' \cdot D' + f' \cdot \mu'_2)^T \end{bmatrix} \cdot w = \begin{bmatrix} \mu_{14} \\ \mu_{24} \\ \mu'_{14} \\ \mu'_{24} \end{bmatrix} \quad (40)$$

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 equations (33) and to recognise the position of a generic point starting to its camera image.

6.1 The stereoscopic problem [2]

Relation (40) represents the stereoscopic problem, it consists in a system of 4 equations in 3 unknown, in the form:

$$A(u, u', v', w) \cdot w = B \quad (41)$$

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.

Starting to first equation of (33), it is possible to write:

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

$$\frac{1}{\delta_u} (\xi_x \cdot w_x + \eta_x \cdot w_y + \zeta_x \cdot w_z) - \quad (42)$$

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

By means of equation (9), it is possible to write:

$$\begin{aligned}
 & D_x \cdot w_x + D_x \cdot w_y + D_x \cdot w_z = \\
 & w_x (\xi_x \cdot n_{r,\xi} + \xi_y \cdot n_{r,\eta} + \xi_z \cdot n_{r,\zeta}) + w_y (\eta_x \cdot n_{r,\xi} + \eta_y \cdot n_{r,\eta} + \eta_z \cdot n_{r,\zeta}) + \\
 & w_z (\zeta_x \cdot n_{r,\xi} + \zeta_y \cdot n_{r,\eta} + \zeta_z \cdot n_{r,\zeta}) + \\
 & (t_\xi \cdot n_{r,\xi} + t_\eta \cdot n_{r,\eta} + t_\zeta \cdot n_{r,\zeta})
 \end{aligned} \quad (43)$$

If we define the elements:

$$\begin{aligned}
 N_\xi &= (\xi_x \cdot n_{r,\xi} + \xi_y \cdot n_{r,\eta} + \xi_z \cdot n_{r,\zeta}); \\
 N_\eta &= (\eta_x \cdot n_{r,\xi} + \eta_y \cdot n_{r,\eta} + \eta_z \cdot n_{r,\zeta}); \\
 N_\zeta &= (\zeta_x \cdot n_{r,\xi} + \zeta_y \cdot n_{r,\eta} + \zeta_z \cdot n_{r,\zeta}); \\
 k &= (t_\xi \cdot n_{r,\xi} + t_\eta \cdot n_{r,\eta} + t_\zeta \cdot n_{r,\zeta}),
 \end{aligned} \quad (44)$$

equation (33) becomes:

$$\begin{aligned}
 & \left(\frac{\xi_x}{\delta_u} - \frac{(u - u_0) \cdot N_\xi}{f} \right) \cdot w_x + \left(\frac{\eta_x}{\delta_u} - \frac{(u - u_0) \cdot N_\eta}{f} \right) \cdot w_y + \left(\frac{\zeta_x}{\delta_u} - \frac{(u - u_0) \cdot N_\xi}{f} \right) \cdot w_z + \\
 & \frac{u - u_0}{f} \cdot k = - \frac{t_\xi}{\delta_u}
 \end{aligned} \quad (45)$$

An analogous relation can be written for second equation of (33):

$$\begin{aligned}
 & \left(\frac{\xi_y}{\delta_v} - \frac{(v - v_0) \cdot N_\xi}{f} \right) \cdot w_x + \left(\frac{\eta_y}{\delta_v} - \frac{(v - v_0) \cdot N_\eta}{f} \right) \cdot w_y + \left(\frac{\zeta_y}{\delta_v} - \frac{(v - v_0) \cdot N_\xi}{f} \right) \cdot w_z + \\
 & \frac{v - v_0}{f} \cdot k = - \frac{t_\eta}{\delta_v}
 \end{aligned} \quad (46)$$

By arranging equation (45) and (46), it is possible to redefine the stereoscopic problem, expressed by equation (40):

$$P(u, u', v') \cdot w = S \quad (47)$$

In equation (47) P is a matrix 4×3 , whose elements depend only by (u, v) and (u', v') , and B is a vector 4×1 , whose elements contain parameters of cameras configuration.

The expression of matrix P is:

$$P = \begin{bmatrix} \frac{\xi_x}{\delta_u} - \frac{(u-u_0) \cdot N_\xi}{f} & \frac{\eta_x}{\delta_u} - \frac{(u-u_0) \cdot N_\eta}{f} & \frac{\zeta_x}{\delta_u} - \frac{(u-u_0) \cdot N_\xi}{f} \\ \frac{\xi_y}{\delta_v} - \frac{(v-v_0) \cdot N_\xi}{f} & \frac{\eta_y}{\delta_v} - \frac{(v-v_0) \cdot N_\eta}{f} & \frac{\zeta_y}{\delta_v} - \frac{(v-v_0) \cdot N_\xi}{f} \\ \frac{\xi'_x}{\delta_{u'}} - \frac{(u'-u'_0) \cdot N_{\xi'}}{f'} & \frac{\eta'_x}{\delta_{u'}} - \frac{(u'-u'_0) \cdot N_{\eta'}}{f'} & \frac{\zeta'_x}{\delta_{u'}} - \frac{(u'-u'_0) \cdot N_{\xi'}}{f'} \\ \frac{\xi'_y}{\delta_{v'}} - \frac{(v'-v'_0) \cdot N_{\xi'}}{f'} & \frac{\eta'_y}{\delta_{v'}} - \frac{(v'-v'_0) \cdot N_{\eta'}}{f'} & \frac{\zeta'_y}{\delta_{v'}} - \frac{(v'-v'_0) \cdot N_{\xi'}}{f'} \\ \frac{\delta_{u'}}{f'} & \frac{\delta_{v'}}{f'} & \frac{\delta_{u'}}{f'} \end{bmatrix} \quad (48)$$

The expression of vector S is:

$$S = \begin{bmatrix} -\frac{t_\xi}{\delta_u} - \frac{u-u_0}{f} \cdot k \\ -\frac{t_\eta}{\delta_v} - \frac{v-v_0}{f} \cdot k \\ -\frac{t_{\xi'}}{\delta_{u'}} - \frac{u'-u'_0}{f'} \cdot k' \\ -\frac{t_{\eta'}}{\delta_{v'}} - \frac{v'-v'_0}{f'} \cdot k' \end{bmatrix} \quad (49)$$

By equation (47) it is possible to invert the problem that is described by eqs. (33) and to recognise the position of a generic point starting to its camera image, by means of pseudoinverse matrix P^+ of matrix P .

$$P \cdot w = S \Rightarrow P^T \cdot P \cdot w = P^T \cdot S \Rightarrow w = (P^T \cdot P)^{-1} \cdot P^T \cdot S \Rightarrow w = P^+ \cdot S \quad (50)$$

By means of equation (50), it is possible to solve the stereoscopic problem in all configurations in which is verified the condition:

7. The camera calibration [2, 18]

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 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.

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. 5.

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).

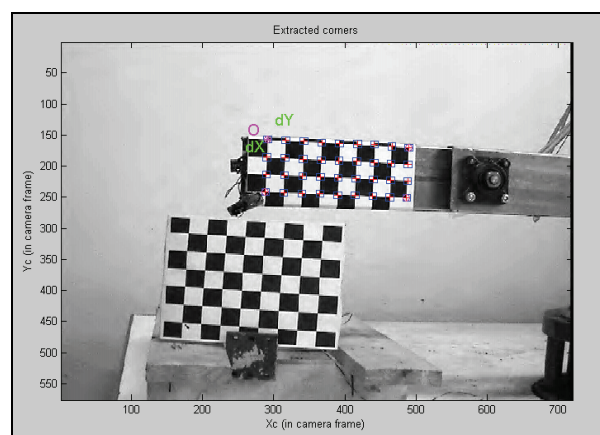


Figure 5. Calibration image

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 $6 \cdot n$ 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. 6.

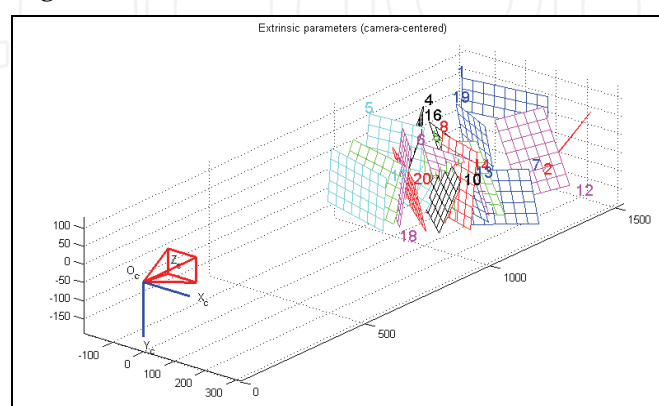


Figure 6. Position of the grids for the calibration procedure

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. 7.

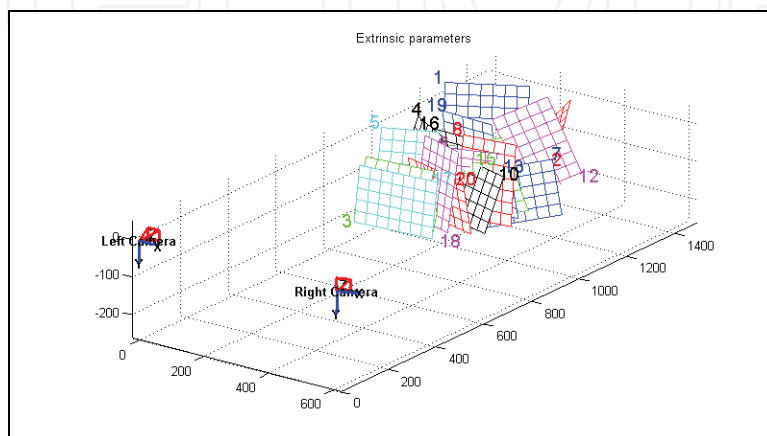


Figure 7. Calibration planes

8. 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 Hartenberg 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.

8.1 The calibration technique [3, 7]

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:

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}, \begin{pmatrix} \theta_{1,i} \\ \theta_{2,i} \\ \theta_{3,i} \end{pmatrix} \quad (51)$$

where: $i = 1, \dots, N$, 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 $i-1$:

$${}^{i-1}A_i = \begin{bmatrix} C\vartheta_1 & -C\alpha_1 \cdot S\vartheta_1 & S\alpha_1 \cdot S\vartheta_1 & a_1 \cdot C\vartheta_1 \\ S\vartheta_1 & C\alpha_1 \cdot C\vartheta_1 & -S\alpha_1 \cdot C\vartheta_1 & a_1 \cdot S\vartheta_1 \\ 0 & S\alpha_1 & C\alpha_1 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (52)$$

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) :

$${}^0T_n = {}^0A_1 \cdot {}^1A_2 \cdot \dots \cdot {}^{n-1}A_n \quad (53)$$

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:

$${}^{II}_{DH} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ d_1 \\ d_2 \\ d_3 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \quad (54)$$

represents the unknown of the kinematic calibration problem.

Said:

$$\Theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \quad (55)$$

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

$$w = t_4(\Pi_{DH}, \Theta) \quad (56)$$

where w is the position vector in the first frame and t_4 is the fourth row of the Denavit-Hartenberg transform matrix. In equation (56) 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 Θ .

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:

$$\begin{cases} t_4(\Pi_{DH}, \Theta^1) = w_1 \\ t_4(\Pi_{DH}, \Theta^2) = w_2 \\ \dots\dots\dots \\ t_4(\Pi_{DH}, \Theta^N) = w_N \end{cases} \quad (57)$$

with $N \geq 9$.

As, for each of the camera images the unknown Denavit-Hartenberg parameters are the same, equations (57) 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 (56), find the solutions that minimize the expression:

$$\int_{D_{\Theta}} |t_4(\Pi_{DH}, \Theta) - w|^2 \cdot d\Theta \quad (58)$$

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

$$\sum_{i=1}^N |t_4(\Pi_{DH}, \Theta^i) - w_i|^2 \quad (59)$$

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.

8.2 Camera model and D-H robotic matrix [7, 8]

Can be useful to include D-H transformation matrix of equation (24), in camera model (33), in this way it is possible to obtain a perspective representation of the robot in an image plane by means joint coordinates.

In homogeneous coordinates, using matrix notation, it is possible to write equation (33):

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = \frac{1}{w_{r,\zeta}} [K] \cdot \begin{Bmatrix} w_{r,\xi} \\ w_{r,\eta} \\ w_{r,\zeta} \\ 1 \end{Bmatrix} \quad (60)$$

where matrix K is:

$$[K] = \begin{bmatrix} -\frac{f}{\delta_u} & 0 & u_0 & 0 \\ 0 & -\frac{f}{\delta_v} & v_0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (61)$$

Considering equation (2), it is possible to write (60) in the frame O,x,y,z , external to images :

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = \frac{1}{D_z \cdot w_z} [K] \cdot [T] \begin{Bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{Bmatrix} \quad (62)$$

Considering equation (9), If we define the vector N:

$$\{N\} = \left\{ \xi_x, \xi_y, \xi_z, t_\xi \right\}^T \quad (63)$$

(62) becomes:

$$\begin{Bmatrix} u \\ v \\ 0 \\ 1 \end{Bmatrix} = \frac{1}{\{N\}^T \cdot \{w\}} [K] \cdot [T] \begin{Bmatrix} w_x \\ w_y \\ w_z \\ 1 \end{Bmatrix} \quad (64)$$

Equation (64) represents the relation between coordinates (u,v) of an assigned point, (e.g. a robot end-effector point expressed in pixels in the image plane) and the coordinates of the same point in the world (Cartesian) frame. In this equation, it is possible to include D-H transformation matrix, to obtain a model that describes the relation between coordinates (u,v) of robot end-effector expressed in pixels, in image plane, and end-effector coordinates in the robot joints space. The relation that synthetizes the model is following:

$$\{u,v\} = \frac{1}{\{N\}^T \cdot [T_n^0] \{\tilde{w}\}_n} [K] \cdot [T] [T_n^0] \{\tilde{w}\}_n \quad (65)$$

where:

- $\{u,v\}$: vector with end-effector coordinates expressed in pixel in image plane;

- $\{\tilde{w}\}_n$: end-effector homogeneous coordinates in robot frame n , for a generic robot with n d.o.f;
- $[T_n^0]$: Denavit-Hartenberg robot transformation matrix from base frame to end-effector frame;
- $[T]$: transformation matrix from camera frame to robot base frame;
- $[K]$: matrix with geometric and optical camera parameters;
- $\{N\}$: vector with expression of optic axis in robot base frame.

8.3 Experimental results

Experimental tests have been executed on a revolute robot prototype with 3 d.o.f., in order to verify the effectiveness of the algorithm.

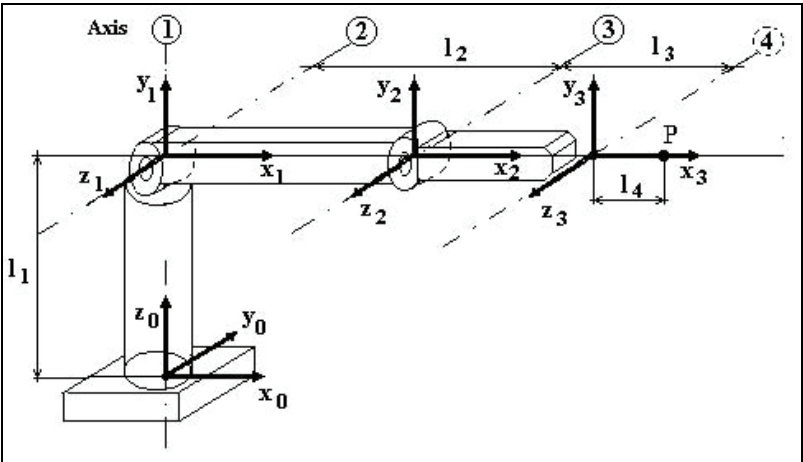


Figure 8. Revolute robot scheme

Twenty images of the robot in twenty different positions of its workspace, with two cameras, have been acquired. After a vision system calibration, by means of an optimization algorithm that uses minimum square technique, it is possible to solve the system of equations (65) and to obtain a numerical solution. Using two cameras we have $2 \cdot 20 = 40$ equations (65), to find nine D-H parameters that characterize the kinematics structure of a three axis revolute robot.

In the tables 1, real parameters (real) and calculated parameters (comp.) are shown.

Joint	a_i (mm)		α_i (deg)		Θ_i (deg)		d_i (mm)	
	Real	Comp.	Real	Comp.	Real	Comp.	Real	Comp.
1	0	7.04mm	90°	86.28°	$-180^\circ \div 180^\circ$	$-180^\circ \div 180^\circ$	$l_1 = 449\text{mm}$	$l_1 = 447.15\text{mm}$
2	$l_2 = 400\text{mm}$	$l_2 = 396.16\text{mm}$	0°	0.94°	$-90^\circ \div 45^\circ$	$-90^\circ \div 45^\circ$	0	10.85mm
3	$l_3 = 400\text{mm}$	$l_3 = 413.65\text{mm}$	0°	0°	$-90^\circ \div 90^\circ$	$-90^\circ \div 90^\circ$	0	13.75mm

Table 1. Real and calculated prototype D-H parameters

9. 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 figure 9 are reported a couple of frames, respectively, from the right telecamera and the left one. In fig. 10 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.

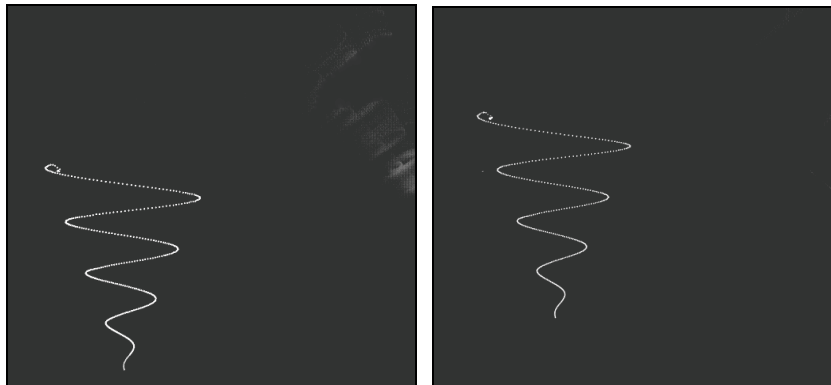


Figure 9. Trajectories in image space: camera position 1(left), camera position 2 (right)

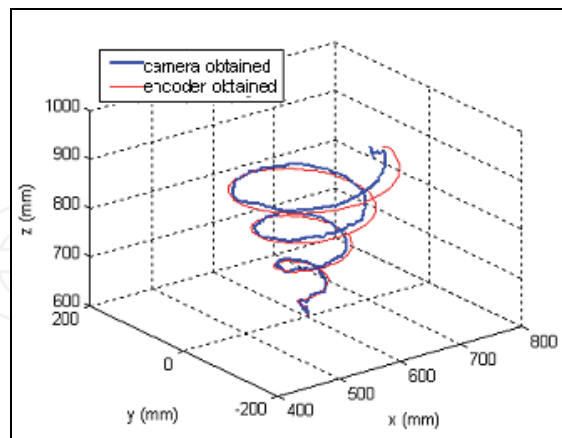


Figure 10. Comparison between trajectory recordings

10. Path planning by means a video system [9]

Was developed a software that allows to choose the end-effector trajectory points. By means of this software, it is possible to select “objective” points, for which the robot must journey, e “obstacle” points, that must be avoided.

The software recognizes the positions of such points in the work space, using a developed camera model.

The procedure starts from a couple of images (taken from two different cameras, fig. 11); the operator selects (with the cursor) a point on the first image of the couple and this will fix a point in a plane. Subsequently, on the second image appears a green line, that represents the straight line that links the focus of the first camera to that point. Now the operator can fix the real position (in the work space) of that point by clicking on this green line.

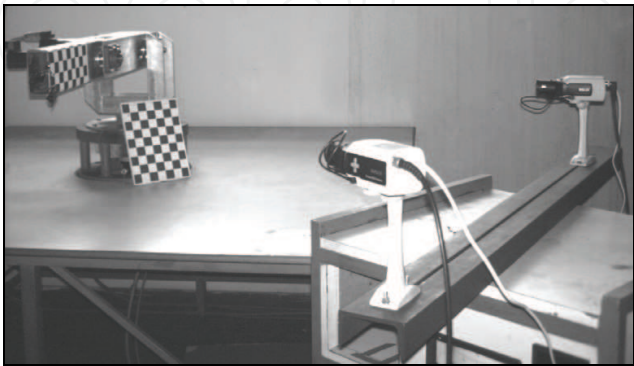


Figure 11. The stereoscopic vision system

In figure 12 the couple of images is reported; on the left is reported the first image and on the right the second one; on the second image is also reported a white solid thick line that is the line that links the focus of the first camera to the point selected on the image on the left.

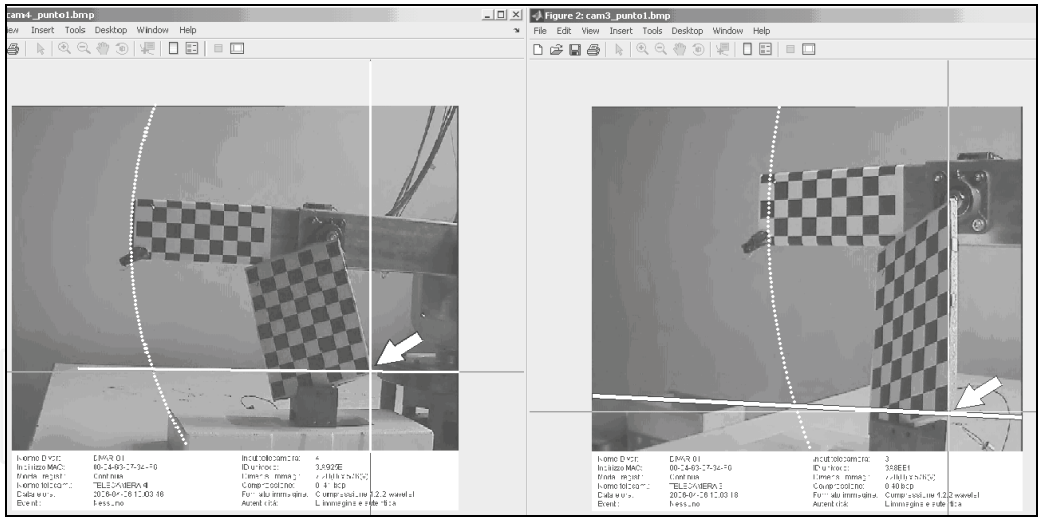


Figure 12. Point assigning by the couple of images

This procedure gives the coordinates of the selected point in the frame of the working space (world frame). Once a point has been assigned in the work space, by means of inverse kinematics it is possible to compute the joint coordinates of the robot when the robot's end-effector is in that position.

Finally the procedure permits to assign a point either as belonging to the path, or as representing an obstacle; in this last case, the path will be computed in order to avoid that point.

In figure 13 the robot arm and the work space are shown; the numbers 1, 2 and 3 represent three points of the path and the cardinals I and II represent two obstacles that are supposed to be spherical.

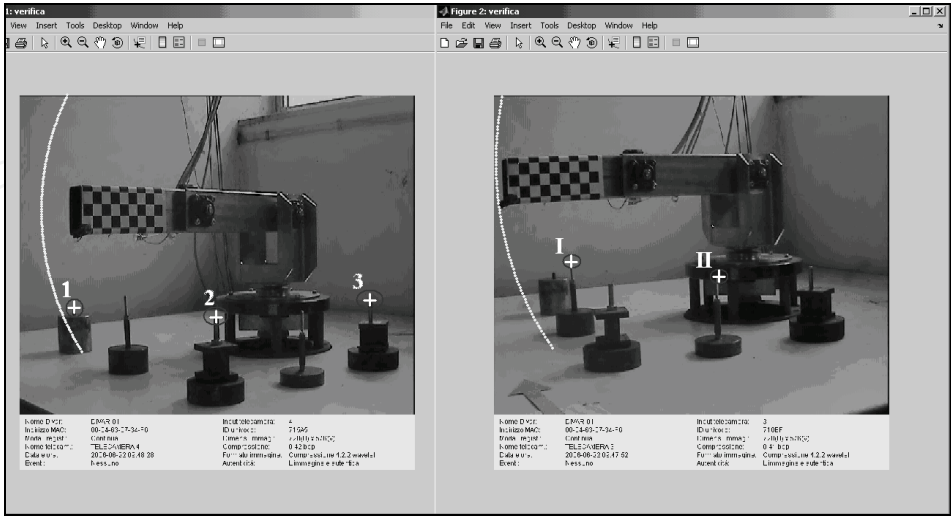


Figure 13. Path assigning

It has to be pointed out that, as previously told, to fix a point a couple of images is needed; in figure 13, for the sake of simplicity, just two images are reposted: on the left is the first image of the couple used for the points, while on the right is reported the second image of the couple for the obstacles.

The path is made up by straight segments that link the selected points (those belonging to the path). To every point that represents an obstacle, is associated the center of a sphere, the sphere radius depends by obstacle dimensions and it is chosen when the procedure starts.

If one of straight segment intersects one of this sphere, the procedure records these intersections and joints each couple of them by means of an arc of a circle. So, the path will consist in a number of straight segments and arcs of circle.

The operator has the possibility to choose the density of the segments and arcs intermediate points, the necessary time to the description of the trajectory, and the obstacles dimensions.

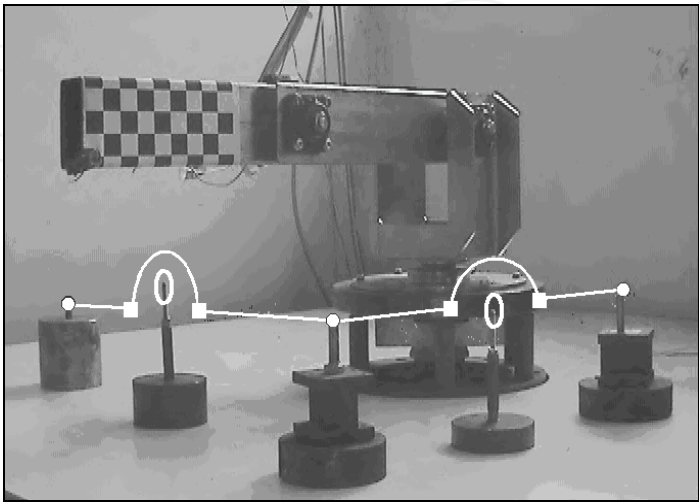


Figure 14. Example of path in the work space

In figure 14 the robot arm and an example of path are shown. In the same figure the points and the obstacles are, also, clearly visible. The points are marked with the same meanings used in the previous figure.

11. Solid reconstruction with a video system on a robot arm

The use of a camera in a robot application, can be performed with two types of architecture: the camera is said eye-in-hand when rigidly mounted on the robot end-effector and it is said eye-to-hand when it observes the robot within its work space. These two schemes have technical differences and they can play very complementary parts. Obviously, the eye-in-hand one has a partial but precise sight of the scene whereas the eye-to-hand camera has a less precise but global sight of it.

Eye-in-hand systems are used primarily to guide robot end effectors and grippers, and to ensure that grippers properly engage the intended targets. The system can also precisely measure the distance from the end effector or gripper to a target. In a robot equipped with an eye-in-hand system, it allows positive identification of a target. In this way it is possible to use the system also to reconstruct a solid in the robot workspace, by means different camera placements and robot inverse cinematic. The next subsection will focus on one commonly used image-based reconstruction methods: *Shape From Silhouettes*.

11.1 A 3D reconstruction technique: Shape From Silhouettes and Space Carving [10, 11, 12]

Shape From Silhouettes is well-known technique for estimating 3D shape from its multiple 2D images.

Intuitively the silhouette is the profile of an object, comprehensive of its inside part. In the “Shape from Silhouette” technique silhouette is defined like a binary image, which value in a certain point (x, y) underlines if the optical ray that passes for the pixel (x, y) intersects or not the object surface in the scene. In this way, Every point of the silhouette, respectively of value “1” or “0”, identifies an optical ray that intersects or not the object.



Figure 15. A computer mouse: the object acquired image (left), the computed object silhouette region (right)

To obtain object volume from silhouettes, we use the space carving technique. A 3D box is modelled to be an initial volume model that contains the object. This box is divided in discrete elements called voxels. The algorithm is performed by projecting the center of each voxel into each image plane, by means of the known intrinsic and extrinsic camera parameters (fig. 16). If the projected point is not contained in the silhouette region, the voxel is removed from the object volume model.

The accuracy of the reconstruction obtained depends on the number of images used, on the positions of each viewpoint considered, on the camera's calibration quality and on the complexity of the object shape.

Using the camera on the robot, is of great aid because, in this way, we know exactly the position of the camera reference frame in the robot work space. Therefore the camera extrinsic parameters, are known without a vision system calibration and it's easy to make an elevated number of photos.

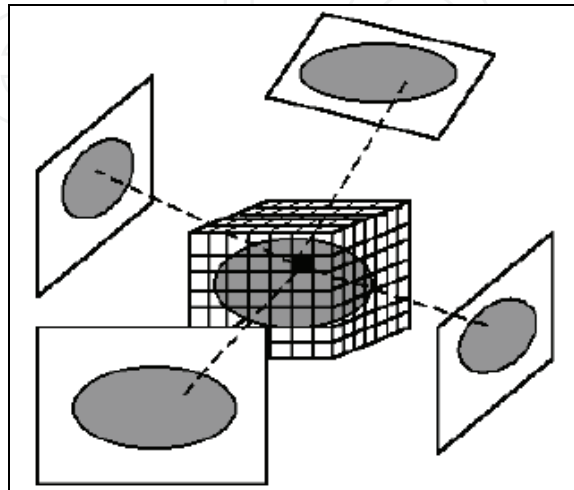


Figure 16. Space carving technique algorithm scheme

One digital camera and a robot prototype, that was designed and built at our laboratory, are used, like show in figure 17. The images have a resolution of 2592×1944 pixels and are saved in raw RGB format.

By means of a turntable, it is possible also to rotate the object, around a vertical axis, of a known angle. These rotations with robot movements allow to capture object images from all its sides and with different angles-shot. In this way, it is possible to use a robot with only three axis, to photograph the objects from all angles-shot.

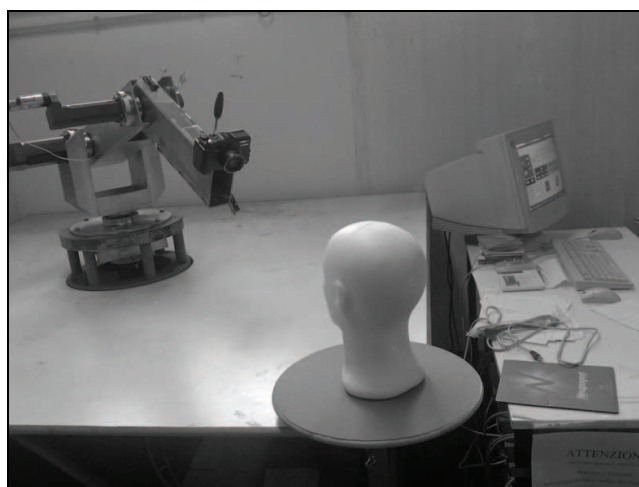


Figure 17. Acquisition system

An algorithm that use this technique was implemented; it ~~and~~ can be divided in three steps:

1. images analysis and object silhouette reconstruction;

2. calculation of the transformation matrix that permits to pass from work space coordinate to each image plane coordinate;
3. 3D-solid reconstruction.

Intersections of the optical axes of camera for each positions with horizontal reference plane of robot reference system $Oxyz$, are evaluated to choose object volume position. Subsequently it is possible to divide the initial volume model in a number of voxels according to the established precision. The centers of voxels are projected into each image plane by means of the pin-hole camera model. In this way it is possible to construct a matrix with the same dimension of image matrix, that has non zero-values only for volume projected voxels. The object silhouette, in the image, is represented by another matrix with non-zero values only for points of silhouette.

The elements of the product among the two matrix that have non-null value are re-transformed in the work space and they became the centers of the voxels that must be used for the following image.

This procedure is repeated for all images, to obtain the volume object in the robot workspace.

11.2 Experimental results

The reconstructions of two objects are presented to demonstrate the performance of the algorithm. As displayed in fig. 18, the test objects are a computer mouse and an mockup head.

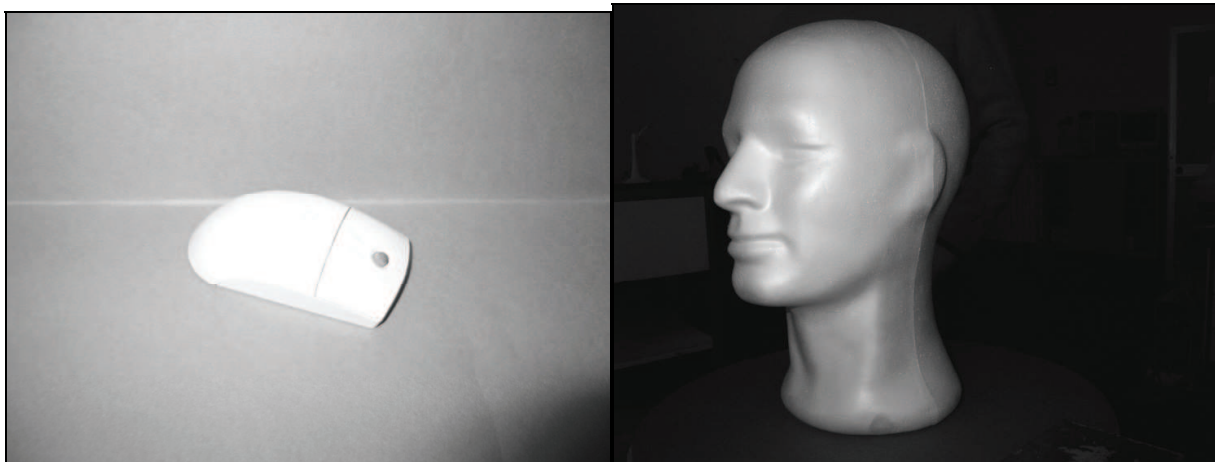


Figure 18. Test objects

Tests are carried out by varying a parameter that represents the resolution of volume.

The resolutions differ in base to the number of voxels in a fixed initial volume, for example to a resolution $res = n$ correspond n^3 initial voxels. Assuming a initial box with sides length x y z , the tolerances corresponding to each edge are x/res , y/res , z/res .

For the mouse reconstruction, 8 photos are been used, while for the reconstruction of the head, 24 photos are been used. This is due to the greater complexity of the head shape regarding the mouse shape.

With the data of the object shape, it is possible to plan robot trajectories to reproduce the object form in any point of the robot workspace, in any position and with any scale factor.

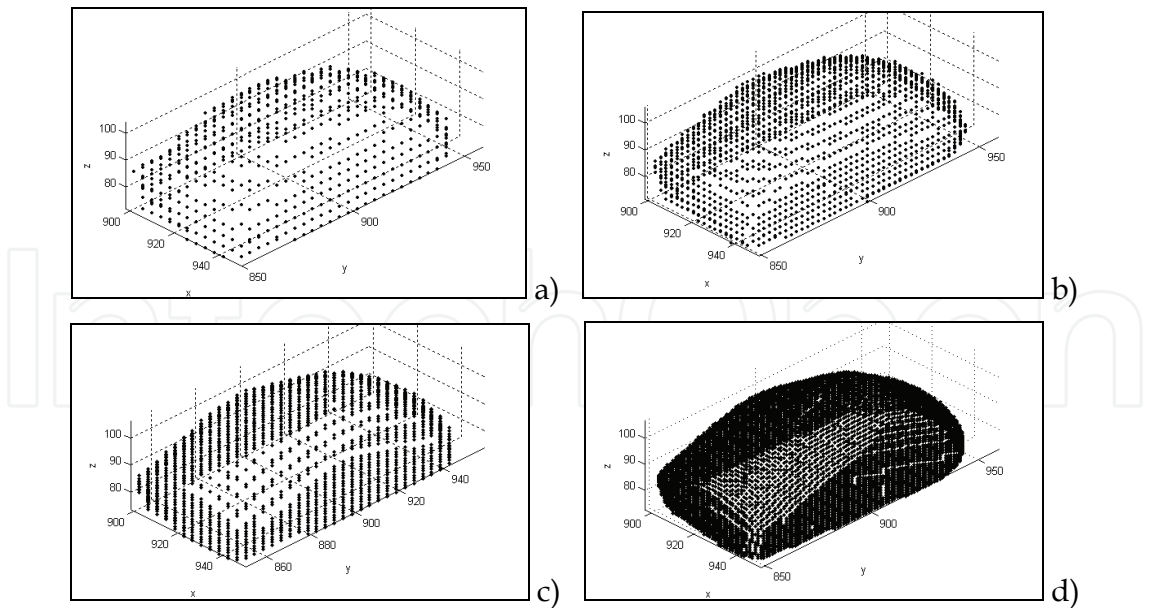


Figure 19. Reconstructed computer mouse: a) res = 50, b) res = 80, c)res=100, d) res = 150

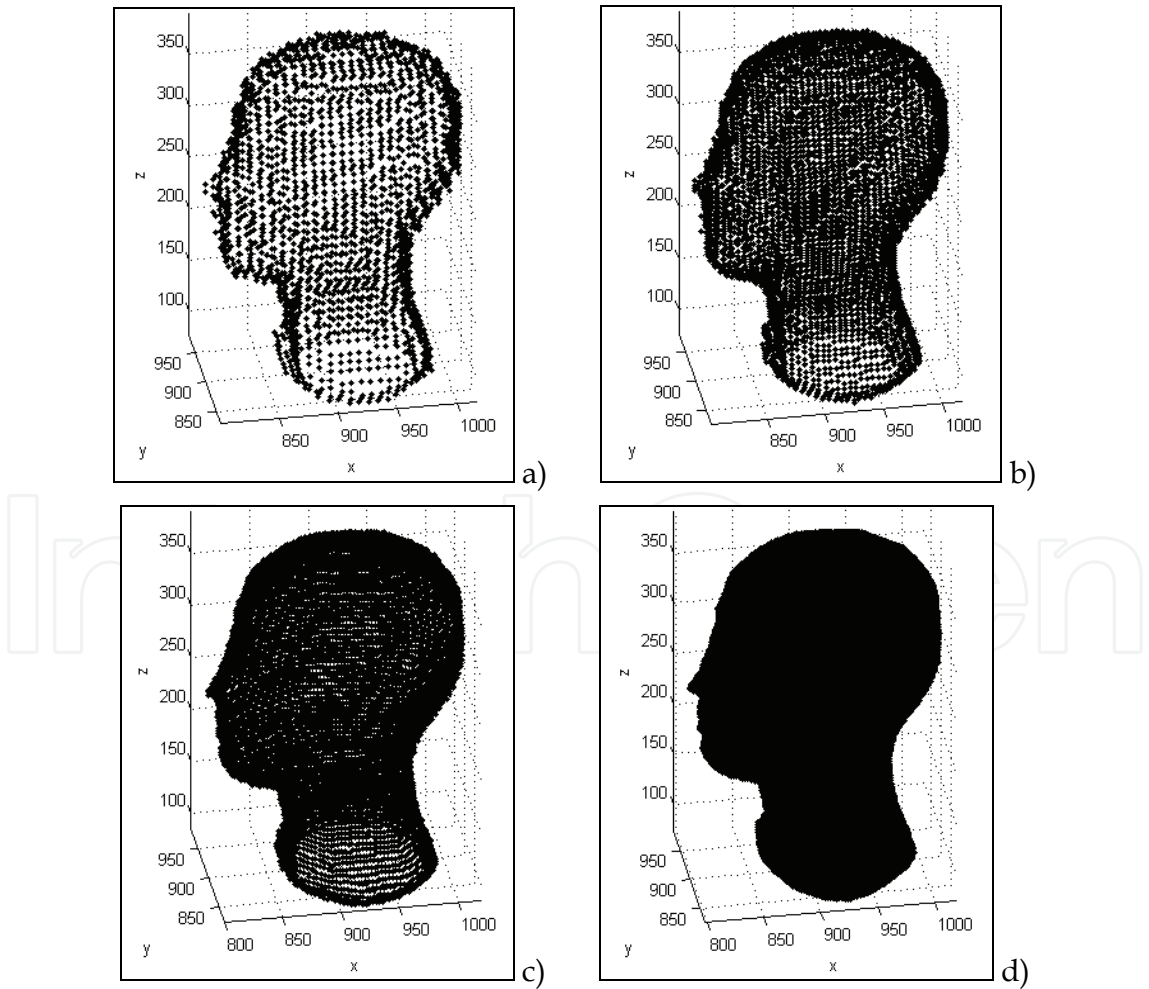


Figure 20. Reconstructed head: a) res = 50, b) res = 80, c) res = 100, d) res = 150

The simplest trajectory that can be plan with information of 3D reconstruction, is a continuous line that is bundled up around reconstructed form, passing for all characterized points. In this way a filament winding process is simulated. An example of this kind of trajectory is shown in figure 21 for a computer mouse and for a mockup head.

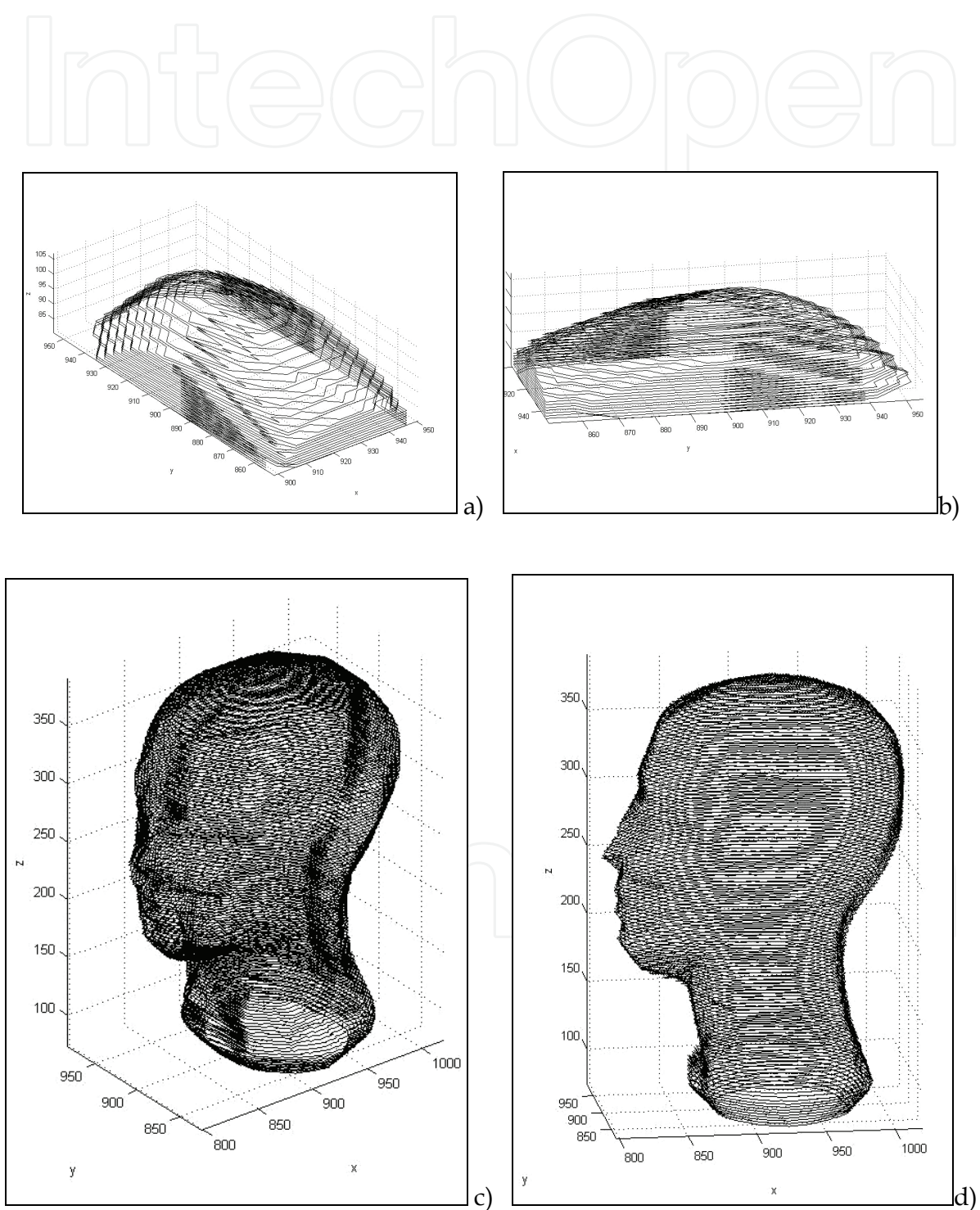


Figure 21. Robot trajectory to reproduce : a),b)computer mouse; c), d) mockup head

12. References

- Niola, V.; Rossi, C. & Savino S. (2006). Perspective Transform and Vision System for Robotic Applications, *Proceedings of 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation*, February 15-17, 2006, Madrid. [1]
- Niola, V.; Rossi, C. & Savino S. (2006). Modelling and Calibration of a Camera for Robot Trajectories Recording, *Proceedings of 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation*, February 15-17, 2006, Madrid. [2]
- Niola, V.; Rossi, C. & Savino S. (2006). A Robot Kinematic Calibration Technique - *Proceedings of 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation*, February 15-17, 2006, Madrid. [3]
- Niola, V.; Rossi, C. & Savino S. (2006). A Camera Model for Robot Trajectories Recording , Published on International review: *WSEAS Transactions on Computers*, Iusse 2, Vol.5, pp.403 – 409, February 2006. [4]
- Niola, V.; Rossi, C. & Savino S. (2006). Perspective Transform in Robotic Applications, Published on International review: *WSEAS Transactions on Systems*, Iusse 4, Vol.5, pp.678 – 684, April 2006,. [5]
- Niola, V.; Rossi, C. & Savino S. (2007). An Application of Vision Systems to the Path Planning of Industrial Robots, *Proceedings of BVAI 2007 2nd International Symposium on Brain, Vision and Artificial Intelligence*, vol. 1, October 10-12, 2007. [6]
- Niola, V.; Pollasto, E.; Rossi, C. & Savino S. (2006). An Algorithm for Kinematics Calibration of Robot Arm, *Proceedings of RAAD'06, 15th International Workshop on Robotics in Alpe-Adria-Danube Region*, June 15-17, 2006, Balatonfured. [7]
- Ciccarelli, V; D'Orsi, G.; Proni, A. & Rossi, C.(2006). Early Experimental Tests on a Vision System for Robot Mechanical Calibration, *Proceedings of RAAD'06, 15th International Workshop on Robotics in Alpe-Adria-Danube Region*, pp. 55-62, June 15-17, 2006, Balatonfured. [8]
- Niola, V.; Rossi, C. & Savino S. (2007). Vision System for Industrial Robots Path Planning, *International Journal of Mechanics and Control*, pp. 35-45, ISSN: 1590-8844. [9]
- Azevedo, T. C. S.; Tavares, J. M. R. S. & Vaz, M. A. P., (2007). 3D object reconstruction from uncalibrated images using a single off-the-shelf camera, *Proceedings of VIP IMAGE , Thematic conference on computational vision and medical image processing*, October 17-19, 2007. [10]
- Chalidabhongse, T.H.; Yimyam, P.& Sirisomboon, P.,(2006). 2D/3D Vision-Based Mango's Feature Extraction and Sorting, *ICARCV 2006 (1-6)*. [11]
- Fremont, V.; Chellali, R., (2004). Turntable-Based 3D Object Reconstruction, *Proceedings of IEEE Conference on Cybernetics and Intelligent Systems*, pp. 1276-1281, Singapore, 2004. [12]
- Fusiello, A., (2005). *Visione Computazionale: appunti delle lezioni*, Informatic Department, University of Verona, 3 March 2005. [13]
- Sharma, R.; Hutchinson, S, (1994). Motion perceptibility and its application to active vision-based servo control, *Technical Report UIUC-BI AI RCV-94-05*, The Beckman Institute, Illinois University. [14]
- Sharma, R., (1994). Active vision for visual servoing: a review, *IEEE Workshop on Visual Servoing: Achivement, Application and Open Problems*, May,1994. [15]

- Sharma, R.; Hutchinson, S, (1995). Optimizing hand/eye configuration for visual-servo system. *IEEE International Conference on Robotics and Automation* , pp. 172-177, 1995. [16]
- Sharma, R.; Hutchinson, S, (1999). On the observability of robot motion under active camera control, *Proc. IEEE International Conference on Robotics and Automation*, May 1999, pp. 162-167. [17]
- Mei, C. Camera Calibration Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc. [18]
- J Feddema, T. , Lee C. S. George, Mitchell O. R., (1991). Weighted selection of image features for resolved rate visual feedback control, *IEEE Trans. Robot. Automat.*, Vol. 7,, pp. 31-47. [19]



Brain, Vision and AI

Edited by Cesare Rossi

ISBN 978-953-7619-04-6

Hard cover, 284 pages

Publisher InTech

Published online 01, August, 2008

Published in print edition August, 2008

The aim of this book is to provide new ideas, original results and practical experiences regarding service robotics. This book provides only a small example of this research activity, but it covers a great deal of what has been done in the field recently. Furthermore, it works as a valuable resource for researchers interested in this field.

How to reference

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

Vincenzo Niola, Cesare Rossi, Sergio Savino and Salvatore Strano (2008). Video System in Robotic Applications, Brain, Vision and AI, Cesare Rossi (Ed.), ISBN: 978-953-7619-04-6, InTech, Available from: http://www.intechopen.com/books/brain_vision_and_ai/video_system_in_robotic_applications

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen