

Singularities in the determination of the situation of a robot effector from the perspective view of 3 points

Henri Michel, Patrick Rives

► **To cite this version:**

Henri Michel, Patrick Rives. Singularities in the determination of the situation of a robot effector from the perspective view of 3 points. [Research Report] RR-1850, INRIA. 1993. <inria-00074822>

HAL Id: inria-00074822

<https://hal.inria.fr/inria-00074822>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNITÉ DE RECHERCHE
INRIA-SOPHIA ANTIPOLIS

Institut National
de Recherche
en Informatique
et en Automatique

2004 route des Lucioles
B.P. 93
06902 Sophia-Antipolis
France

Rapports de Recherche

N°1850

Programme 4

Robotique, Image et Vision

SINGULARITIES IN THE DETERMINATION OF THE SITUATION OF A ROBOT EFFECTOR FROM THE PERSPECTIVE VIEW OF 3 POINTS

Henri MICHEL
Patrick RIVES

Fevrier 1993

Singularities in the determination of the situation of a robot effector from the perspective view of 3 points

Henri Michel, Patrick Rives

INRIA Sophia Antipolis, 2004 Route des Lucioles, 06902 SOPHIA ANTIPOLIS Cedex

Abstract

With the advent of new visual sensors technology and improved image processing methods, the possibility of relying only on visual information to design closed-loop commands for robot control is not an utopia anymore.

This approach brings many advantages such as direct use of image characteristics in control. Instead of intervening at different stages in the control process, robot controllers and image processing modules closely cooperate to achieve a vast collection of various robotics tasks such as assembling mechanical parts or driving a mobile robot.

This concept: *visual servoing*, though fruitful, leads us to face new and unsolved issues. Among them, one is crucial for practical design of such vision-based systems. The information extracted from the images provided by the cameras must be incorporated in a controllable scheme for which a robust command must be found. These control theory considerations can be expressed as: the mapping between the space where controls are performed and the space where measures are observed must be *regular* enough. The most obvious of these *regularity conditions* is: we should not encounter singularities. What we call singularities are configurations where the rank of the differential of these mapping functions decreases. In practice, those singularities would account for cases where we should have to solve ill conditioned problems and find an inverse for singular matrices.

We give here a complete solution of such a problem when the space of controls is the 6 degrees of freedom of a robot end-effector (3 positions plus 3 angles) and the space of measurements is the set of triplets of projected points. These points are the perspective projections counterparts in 2D space of known triplets of rigidly connected points in 3D space. In practical realizations these points could be, for example, centers of circles determined from a CAD model of a mechanical part, or in an automated cartography context, landmarks on the ground.

Keywords: Robot Control, sensor based control, Visual servoing, singularities of perspective projection, PnP problem.

1 Introduction

The concept of *visual servoing* was introduced recently (see [13] for a reference). The originality of this approach was to describe a robotics task in terms of features in an image. The task is achieved when what we observe in a camera attached to a robot or vehicle fits a precomputed reference image. For example let us consider the task of having a mobile robot pass through a door. A classical method to achieve this task is to give a set of positions describing an acceptable path through the door, then perform a control on the difference between those reference positions and the actual ones. A visual servoing version would be to compute a set of reference views of this door while we cross it, then perform a control based on the differences of these precomputed views and what is actually observed in the camera image.

This approach is based on a strong underlying formalism based on two classic robotics concepts. First, we use a description of a task in terms of *links* similar to classical mechanical links. Imposing a certain view of a know object can be seen as attaching a *link* between a coordinate frame on this object and the camera coordinate frame. Second, the practical design of a closed loop control to achieve this task is derived from the classical task function framework.

The following section will show more precisely how to describe image characteristics for a visual servoing scheme, then how to derive a control law suitable for this vision based system referring to the task function approach. We will notice that the combination of typical robotics considerations with vision specific issues lead to superpose robot singularities to vision singularities, though these problems are formally similar. Our new result on the singularities of the peculiar 3 points problem will be related to the classical 3 points *local determination* problem presented in [4].

2 Problem Statement

2.1 Visual Signals

The choice of image characteristics that are needed to perform a visual servoing task depends on a blend of image recognition and control criteria as described in [3]. Among the image-recognition criteria, the cost of feature extraction often restricts the choice of visual characteristics to “simple” image features such as points, lines segments or more complex algebraic curves that can be extracted at video rate. Some authors have used features related to regions instead of curves in the image such as surface area, center of gravity and inertia axis for objects tracking [5], or positioning [13]. Other methods have been investigated, for example using phase shift in the Fourier transform of images as described in [7]. A very original approach appears in [12] where the image is described in terms of a graph whose leaves are surfaces in the image and arcs are vertices.

For our purpose we will assume that the image can be described by a vector of *visual signals*. Examples of such visual signals are the coordinates of lighted pixels in the image, or the coefficients of several algebraic curves that fit the set of lighted pixels in the image.

Our main assumption is that what we see in the image only depends on the relative situation between the 3D scene and the camera. Changes in scene illumination, for example, are not taken into account. We also suppose that the 3D scene includes only a finite number

of rigid objects.

So, we deal with visual signals functions of the form :

$$s(t) = s(F_{\text{effector}}^{\text{object}_1}(t) \dots F_{\text{effector}_n}^{\text{object}_n}(t))$$

- s is the representation of an image feature in terms of a vector of visual signals that are functions of time.
- $F_{\text{effector}_i}^{\text{object}_i}$ represents the situation of a coordinate frame attached to a rigid object i with respect to a coordinate frame attached to the camera. Its expression is $R_{\text{object}_i} R_{\text{effector}_i}^{-1}$, R_α is the homogeneous matrix which describes a coordinate frame attached to element α in a reference frame. We will generally call R_α the “*situation*” of α .

In what follows we will generally use capital letters to refer to entities in the 3D scene, and lower case letters for entities in the image.

Our representation of the camera will be the classical pin hole model with focal length equal to 1. A 3D point M of coordinates $(X Y Z)$ in a suitable frame attached to the camera will be projected on the image plane on a 2D point m of coordinates $(x y)$, as following :

$$x = \frac{X}{Z}, y = \frac{Y}{Z} \quad (1)$$

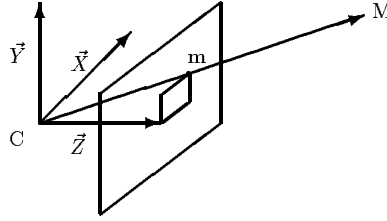


Figure 1: Model of the camera

Any other representation based on projective geometry involving focal length, x-axis scaling factor, offset between the position of the focal axis and the origin of the coordinate frame in the image . . . can be represented by a pin hole model up to two affine transformations in both image and 3D scene coordinates. Thus, such representation is not more general for our purpose since an affine transformation is continuous at any order of differentiation. In what follows we will call “*the camera coordinate frame*” an arbitrary frame, attached to the camera, in which equation 1 holds.

We can represent the motion of the camera coordinate frame by a velocity screw $T = (V_c \Omega_c)$, V_c being the velocity of the optical center of the camera and Ω_c the rotation velocity of the coordinate frame attached to the camera. The velocity in the camera coordinate frame of a static point in 3D scene is :

$$\frac{d \vec{CM}}{dt} = -V_c - \Omega_c \times \vec{CM} \quad (2)$$

If we differentiate equation (1) and use the expression (2), we obtain the classical *equation of optic flow* :

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = - \begin{pmatrix} \frac{1}{Z} & 0 & -\frac{x}{Z} & -xy & 1+x^2 & -y \\ 0 & \frac{1}{Z} & -\frac{y}{Z} & -(1+y^2) & xy & x \end{pmatrix} \cdot T \quad (3)$$

This equation relates the motion of the camera to the apparent motion of a point of the 3D scene observed in the image. This equation is in fact the most trivial example of a relation between a collection of 2D information indexed by time (the visual signals through time) and the collection of 3D situations (3 positions and 3 angles) of the camera. This type of relation can be generalized to a larger class of visual signals using the framework presented in the next section.

2.2 The task function approach

A now classical approach in robotics is to consider the process of achieving a task such as tracking or positioning as a problem of regulation to zero of a certain function: *the task function*. For example, if the imposed task is to follow a given trajectory in cartesian space, the task function approach version of this problem would be: keep $\|e(t)\|$ minimum through time. Where e *the task function vector* is: $e(t) = \bar{R}(q(t)) - R^*(t)$, R denotes the situation of a frame attached to the end effector expressed in a reference frame, R^* is the desired trajectory. The control vector in this problem is the set of joint torques we apply to change the joint coordinates (denoted q in our equations). The task is achieved when $\|e(t)\| = 0$ (see [11] for full details about the task function approach in robot control). The application of the task function approach to sensor based tasks is straightforward: the task is now described as a certain set of image features to obtain. In mathematical terms the general definition of a visual task function vector e is :

$$e(t) = C(s(R_{\text{camera}}^{\bar{}}(q(t)), t) - s^*(t)) \quad (4)$$

where

- s^* can be considered as a reference image target to be reached in the image.
- s is the value of the visual features currently observed by the camera. It depends only on the *relative* position between the camera and the scene, hence on the situation of the camera, and on the situation of mobile objects in the 3D scene. This dependence on mobile objects situations is expressed globally as a dependence on time t .
- C is a matrix which allows, for robustness issues, to take into account more visual features than necessary.
- $R_{\text{camera}}^{\bar{}}$ is a representation of the situation of the camera. We will often call it the cartesian coordinates of the robot.
- q are the robot joint coordinates.

The robot coordinates q appear in this equation since we do not have direct control on the cartesian coordinates \bar{R} but rather on the joint coordinates q .

Once again we should emphasize the important concept below the formalism contained in this equation: the task is not described in the cartesian or joint coordinates of a robot

but in terms of image characteristics. So the vision process does not only provide a mere tool to verify that a closed loop based on robot coordinates behaves correctly and achieves the task, but rather permits us to now deal with a closed loop based on vision information.

The quantity which arises naturally in our output regulation problem is \dot{e} , the derivate of the vision task function versus time. Its expression is :

$$\dot{e} = \frac{\partial e}{\partial s} \cdot \frac{\partial s}{\partial \bar{R}} \cdot \frac{\partial \bar{R}}{\partial q} \cdot \dot{q} + \frac{\partial e}{\partial s} \cdot \frac{\partial s}{\partial t} - \dot{s}^* \quad (5)$$

\dot{e} is thus of the form :

$$\dot{e} = A\dot{q} + B$$

The matrix A governs the controllability, robustness and convergence of a feedback control law in visual signals space.

It is important to notice that this matrix is the product of 3 terms which are very different in nature :

- $\frac{\partial \bar{R}}{\partial q}$ is the classical robot Jacobian which appears in virtually all robotics problems.
- $\frac{\partial s}{\partial \bar{R}}$ is what we called the *interaction matrix* or task function Jacobian. It is canonically related to the vision problem without any reference to a particular robot geometric or dynamics constraints whose influence is totally contained in the robot Jacobian.
- $\frac{\partial e}{\partial s}$ depends on the relative importance of each visual signal we take into account. Here arises the tricky problem of visual signals selection, vastly discussed in [3]. Without addressing the details of this selection process, and considering only the criterion of image features extraction robustness, it is easy to realize that image features which are less reliable due to image processing considerations should have less importance in the combination.

A necessary condition for the visual signals output regulation problem (controllability condition) is that the product of the 3 matrices mentioned above has a determinant different from zero (or is full-rank if it is not a square matrix). The characteristics and issues related to the robot Jacobian are well known (see [10] for a reference). The “vision-part” of the problem is: when is the interaction matrix singular? We address this problem for the particular case where the image features are the projection of the 3 rigidly connected points in 3D space.

The interaction matrix generalizes the optic flow equation for differentiable image features. An analytical expression for the interaction matrix when the image features are general algebraic curves can be obtained. This is a consequence of the implicit function theorem: details can be found in [1] and [2]. Unfortunately the derivation of these interaction matrices is rather tedious to say the least.

3 Solution of the task function singularities problem

3.1 Classical Approach

Following what was said in the previous section, we can give the following definition of the singularities of the vision task function :

Definition 1 *The vision task function singularities are the configurations where the rank of the interaction matrix decreases.*

In our particular case (triplet of points) the interaction matrix is just the conjunction of 3 optic flow equations (one for each point) :

$$L^T = \begin{pmatrix} -\frac{1}{Z_1} & 0 & \frac{x_1}{Z_1} & x_1 y_1 & -(1+x_1^2) & y_1 \\ 0 & -\frac{1}{Z_1} & \frac{y_1}{Z_1} & 1+y_1^2 & -x_1 y_1 & -x_1 \\ -\frac{1}{Z_2} & 0 & \frac{x_2}{Z_2} & x_2 y_2 & -(1+x_2^2) & y_2 \\ 0 & -\frac{1}{Z_2} & \frac{y_2}{Z_2} & 1+y_2^2 & -x_2 y_2 & -x_2 \\ -\frac{1}{Z_3} & 0 & \frac{x_3}{Z_3} & x_3 y_3 & -(1+x_3^2) & y_3 \\ 0 & -\frac{1}{Z_3} & \frac{y_3}{Z_3} & 1+y_3^2 & -x_3 y_3 & -x_3 \end{pmatrix} \quad (6)$$

This matrix is generally of rank 6. The classical (and brute force) approach to solve the problem of singularities is to compute the determinant of the matrix L^T and find the roots of the obtained polynomial. A difficulty arises since L^T depends on the x_i and y_i (which are measurable quantities in the image) and on the Z_i which cannot be obtained directly with a single camera. The first idea which comes to mind is to find the Z_i using some a priori knowledge on the 3D scene and then compute the determinant. Unfortunately, it is necessary to solve the *local determination problem* in order to recover the Z_i . This problem leads to equations relating the x_i and y_i to the Z_i with parameters the known distances between the 3 points in 3D space. So, one may think that the singularities problem is even more difficult than the local determination problem. Some results were obtained ([1]) but only for some very peculiar cases for which it is not necessary to solve the underlying 3 points local determination problem. One such case is when all 3 points lie in a plane perpendicular to the optical axis of the camera.

3.2 Equivalent Approach

The intuitive idea is to avoid the classical representation of the motion between the camera and the 3-D object by a kinematic screw but rather to consider the motion of 3 points forming a rigid body in the following more abstract way. Doing so, we just go back to the first definition of the motion of a rigid body.

Definition 2 *the motion of 3 points forming a rigid body is a motion in the 9 dimensional space formed by the 3-D coordinates of the 3 points constrained to remain on the manifold H where the distances between the 3 couples of 2 different points in the triplet are constant.*

We should keep in mind that this manifold H is of dimension 6 if and only if the 3 points are affine independent (not collinear).

3.2.1 Definition of some coordinates systems

So we introduce the following P and D changes of coordinates systems.

The “projective” coordinates We define the transformation P from \mathbb{R}^9 to \mathbb{R}^9 by :

$$\begin{pmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ X_3 \\ Y_3 \\ Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \xrightarrow{P} \begin{pmatrix} x_1 = \frac{X_1}{Z_1} \\ y_1 = \frac{Y_1}{Z_1} \\ x_2 = \frac{X_2}{Z_2} \\ y_2 = \frac{Y_2}{Z_2} \\ x_3 = \frac{X_3}{Z_3} \\ y_3 = \frac{Y_3}{Z_3} \\ Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \quad (7)$$

In other words, we transform the X and Y coordinates into their projective counterparts but keep the Z coordinates constant.

The “ D ” coordinates This transformation will not change the first 6 coordinates but will add the 3 Euclidean distances between all couples of 3-D points. Using the following notation :

$$\bar{X}_i = Z_i \times \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \quad \bar{x}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \quad (8)$$

we will define $D_1 = d(\bar{X}_2, \bar{X}_3), \dots$ with $d(A, B)$ representing the Euclidean distance between the points A and B . The D coordinates can now be defined with :

$$\begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \\ Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \xrightarrow{D} \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \\ D_1 = d(\bar{X}_2, \bar{X}_3) \\ D_2 = d(\bar{X}_1, \bar{X}_3) \\ D_3 = d(\bar{X}_1, \bar{X}_2) \end{pmatrix} \quad (9)$$

3.2.2 characterization of the singularities

We will now consider the transformation T :

$$T = D \circ P \quad (10)$$

where \circ denotes the product of 2 transformations. Given a triplet of points we get the representation of their projections (the 6 first coordinates) plus a generic description of the rigid triplet by the set of 3 distances (the last 3 coordinates). In this framework we can state the following theorem :

Theorem 1 *A necessary and sufficient condition for a configuration to be a singularity of the 3 points inverse perspective transformation is: “ $\det(\text{Jac}(T))$ is null at this configuration”.*

Proof 1 Let us consider the following transformation f from \mathbb{R}^9 to \mathbb{R}^6 :

$$(\bar{X}_1 \bar{X}_2 \bar{X}_3) \xrightarrow{f} (\bar{x}_1 \bar{x}_2 \bar{x}_3)$$

f gives the projections of 3 points of \mathbb{R}^3 . This f transformation is just an adaptation of the P transformation in which we “forget” the 3 last coordinates. The s visual signal function related to our problem is the product of two transformations ($s = f_H \circ \rho$). f_H is the restriction of f to the hypersurface H defined by assigning 3 constant values to the distances between 2 different points. ρ gives the positions of the 3 reference points in 3D space as a function of the situation (6 degrees of freedom) of the rigid object to which they belong. ρ is not singular (and H a 6 dimensional manifold) if and only if the 3 reference points are not collinear. We will split the proof in two cases depending on whether or not the 3 points are collinear.

1. Let us suppose that the 3 points are not collinear. From the computation of the differential of P (which will be done in a following paragraph), it is easy to see that the differential of f is always of rank 6. The image of $d\rho$ (the differential of ρ) is the 6 dimensional space tangent to H , since the 3 points are not collinear. The image of ds ($Im(ds)$) is the image of the restriction of df to this 6 dimensional space. We have the following equivalences :

- the configuration is not singular.
- $\iff Im(ds)$ is 6 dimensional.
- $\iff Ker(df) \cap \text{space tangent to } H = \{\vec{0}\}$.
- $\iff Ker(df)^\perp + \text{“space normal to } H\text{”} = \mathbb{R}^9$.
- $\iff [dx_1 dy_1 dx_2 dy_2 dx_3 dy_3]^\perp + [dD_1 dD_2 dD_3] = \mathbb{R}^9$.
- we note $[V_1 \dots V_n]$ the vector space generated by the vectors $V_1 \dots V_n$.
- $\iff (dx_1 dy_1 dx_2 dy_2 dx_3 dy_3 dD_1 dD_2 dD_3)$ is a set of linearly independent vectors.
- \iff The Jacobian matrix of T is not singular since the above mentioned vectors are the rows of this matrix.

2. Let us suppose that the 3 points are collinear. In this case the ρ transformation is singular since a rotation whose axis is the line containing the 3 points and no sliding velocity leaves the points unchanged in \mathbb{R}^3 . ρ is singular because the 3 equations on distances defining H are not independent for aligned points. $(dD_1 dD_2 dD_3)$ are not linearly independent vectors. Since they form the 3 last rows of the Jacobian matrix of T , this matrix is singular in this case.

3.2.3 Singular configurations for the T transformation

the Jacobian matrix of the T transformation can be easily computed since :

$$Jac(T)_{p0} = Jac(D)_{Pp0} \circ Jac(P)_{p0} \quad (11)$$

And we get the following relation with the determinants of the Jacobian matrices :

$$\det(Jac(T)_{p0}) = \det(Jac(D)_{P(p0)}) \times \det(Jac(P)_{p0}) \quad (12)$$

Where $\det(B)$ stands for the determinant of the B matrix and $Jac(f)$ stands for the Jacobian matrix of the transformation f .

Determinant of the Jacobian matrix of P

$$\text{Jac}(P) = \begin{pmatrix} \frac{1}{Z_1} & 0 & 0 & 0 & 0 & 0 & \frac{-x_1}{Z_1} & 0 & 0 \\ 0 & \frac{1}{Z_1} & 0 & 0 & 0 & 0 & \frac{-y_1}{Z_1} & 0 & 0 \\ 0 & 0 & \frac{1}{Z_2} & 0 & 0 & 0 & 0 & \frac{-x_2}{Z_2} & 0 \\ 0 & 0 & 0 & \frac{1}{Z_2} & 0 & 0 & 0 & \frac{-y_2}{Z_2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{Z_3} & 0 & 0 & 0 & \frac{-x_3}{Z_3} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{Z_3} & 0 & 0 & \frac{-y_3}{Z_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (13)$$

The determinant of this Jacobian is immediately obtained :

$$\det(P) = \frac{1}{Z_1^2 \times Z_2^2 \times Z_3^2} \quad (14)$$

Determinant of the Jacobian matrix of D

$$\text{Jac}(D) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{\partial D_1}{\partial x_1} & \frac{\partial D_1}{\partial y_1} & \frac{\partial D_1}{\partial x_2} & \frac{\partial D_1}{\partial y_2} & \frac{\partial D_1}{\partial x_3} & \frac{\partial D_1}{\partial y_3} & \frac{\partial D_1}{\partial Z_1} & \frac{\partial D_1}{\partial Z_2} & \frac{\partial D_1}{\partial Z_3} \\ \frac{\partial D_2}{\partial x_1} & \frac{\partial D_2}{\partial y_1} & \frac{\partial D_2}{\partial x_2} & \frac{\partial D_2}{\partial y_2} & \frac{\partial D_2}{\partial x_3} & \frac{\partial D_2}{\partial y_3} & \frac{\partial D_2}{\partial Z_1} & \frac{\partial D_2}{\partial Z_2} & \frac{\partial D_2}{\partial Z_3} \\ \frac{\partial D_3}{\partial x_1} & \frac{\partial D_3}{\partial y_1} & \frac{\partial D_3}{\partial x_2} & \frac{\partial D_3}{\partial y_2} & \frac{\partial D_3}{\partial x_3} & \frac{\partial D_3}{\partial y_3} & \frac{\partial D_3}{\partial Z_1} & \frac{\partial D_3}{\partial Z_2} & \frac{\partial D_3}{\partial Z_3} \end{pmatrix} \quad (15)$$

The determinant of this Jacobian matrix is :

$$\det(D) = \det \begin{pmatrix} \frac{\partial D_1}{\partial Z_1} & \frac{\partial D_1}{\partial Z_2} & \frac{\partial D_1}{\partial Z_3} \\ \frac{\partial D_2}{\partial Z_1} & \frac{\partial D_2}{\partial Z_2} & \frac{\partial D_2}{\partial Z_3} \\ \frac{\partial D_3}{\partial Z_1} & \frac{\partial D_3}{\partial Z_2} & \frac{\partial D_3}{\partial Z_3} \end{pmatrix} \quad (16)$$

Given two points M_i of coordinates (X_i, Y_i, Z_i) and M_j , the partial derivative of the Euclidean distance between those 2 points versus a coordinate Z_l is :

$$\frac{\partial D(M_i, M_j)}{\partial Z_l} = \frac{\partial M_i \vec{M}_j}{\partial Z_l} \cdot M_i \vec{M}_j \frac{1}{D(M_i, M_j)}$$

Grouping all previous intermediate results we get with the definition of the matrix \tilde{M} :

$$\tilde{M} = \begin{pmatrix} 0 & C\vec{M}_2 \cdot M_2\vec{M}_3 & C\vec{M}_3 \cdot M_3\vec{M}_2 \\ C\vec{M}_1 \cdot M_3\vec{M}_1 & 0 & -C\vec{M}_3 \cdot M_1\vec{M}_3 \\ C\vec{M}_1 \cdot M_2\vec{M}_1 & -C\vec{M}_2 \cdot M_1\vec{M}_2 & 0 \end{pmatrix} \quad (17)$$

$$\det(P) = \prod_{i=1}^3 \frac{1}{Z_i^3} \times \prod_{i < j} \frac{1}{\|\vec{M}_i - \vec{M}_j\|} \times \det(\tilde{M}) \quad (18)$$

With the notation: C is the center of the camera coordinate frame (“optical center”), thus $C\vec{M}_i$ is the vector of coordinates (X_i, Y_i, Z_i) . So, excluding two trivial cases the singularities of the 3 points inverse perspective problem are the configurations where $\det(\vec{M})$ is zero. These two trivial cases are :

- the 3 points are not all distinct. This is a singularity since the optic flow equations of a certain point intervene at least twice in the interaction matrix.
- one of the Z_i is zero. In which case the perspective projection is undefined.

3.3 Solution of the equation $\det(\vec{M}) = 0$

As can be seen after performing a permutation on rows and columns (which does not change the value of the determinant), the mentioned equation is equivalent to $\det(\tilde{C}) = 0$. \tilde{C} is the characteristic matrix defined by :

Definition 3 We call characteristic matrix of a configuration of the triplet of points $(M_1 M_2 M_3)$ in a frame with origin C the matrix \tilde{C} whose term at row i and column j is :

$$\tilde{C}_{i,j} = C\vec{M}_i \cdot M_j\vec{M}_i \quad (19)$$

We must now state the following remark: instead of solving this equation in the coordinates of the points M_i in a fixed camera frame, it is much better to solve this equation in the coordinates of C with M_i having fixed coordinates. A good choice would be a coordinate frame which has an axis orthogonal to the plane that contains the 3 rigidly connected points. In this coordinate frame, the 3 points have coordinates $(X_1, Y_1, 0)$ $(X_2, Y_2, 0)$ $(X_3, Y_3, 0)$ and C has coordinates (X, Y, Z) . The crucial advantage of this representation is that the requirement “the points are rigidly connected, forming a triangle with fixed lengths” is naturally fulfilled. The equation on the determinant of the characteristic matrix leads to a polynomial equation $P_1 = 0$ where P_1 is a polynomial in $X_1, Y_1, X_2, Y_2, X_3, Y_3, X, Y, Z$.

Remarks : Because each scalar product intervening in the characteristic matrix is of the form $(X_i - X)(X_i - X_j) + (Y_i - Y)(Y_i - Y_j)$, we can state the following remarks.

- P_1 (considered as a polynomial in $X_1, Y_1, X_2, Y_2, X_3, Y_3, X, Y, Z$) is a sum of products of three terms of degree 2. So it is an homogeneous polynomial of degree 6 or less.
- Because of our peculiar choice of a coordinate frame, the vectors $M_j\vec{M}_i$ all have a null z coordinate, P_1 is in fact a polynomial in $X_1, Y_1, X_2, Y_2, X_3, Y_3, X, Y$ only, there is no more Z dependency.
- P_1 considered as a polynomial in X_i, Y_i has a degree less or equal to 6. Considered as a polynomial in X, Y , its degree is less or equal to 3.

3.3.1 Factorization of P_1

Once these preliminary remarks on intervening polynomials stated, we can obtain a factorization of P_1 with no computation. This is based on the following theorem from algebraic geometry :

Theorem 2 *Let X be an hypersurface of reduced equation $f = 0$ (that is to say that f is a prime polynomial) and g a polynomial in several unknowns with complex coefficients which is zero on X , then g is a multiple of f .*

We shall seek some obvious singular configurations. These configurations correspond to points where P_1 will be zero according to theorem 1 proved above. The equations of each singular configurations class will then provide a factorization of P_1 according to theorem 2.

A first class of singular configurations was already mentioned in the proof of theorem 1. It corresponds to cases where the 3 points belong to the same straight line (the triangle is degenerated into a segment). In these cases there exists a rotation along this segment that leaves the image unchanged. So, if the X_i, Y_i are such that the 3 points of coordinates $(X_1, Y_1, 0) \dots (X_3, Y_3, 0)$ are aligned then P_1 is null (see figure 2).

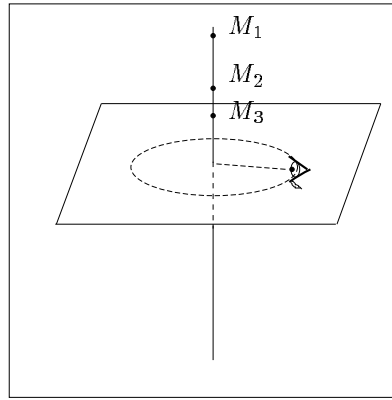


Figure 2: Aligned reference points singularity

The condition so that these 3 points are aligned is formulated as the following polynomial equation :

$$S = \begin{vmatrix} X_2 - X_1 & X_3 - X_1 \\ Y_2 - Y_1 & Y_3 - Y_1 \end{vmatrix} = 0 \quad (20)$$

we use the notation $|M|$ for the determinant of the M matrix. It is easy to see that the absolute value of S (S stands for shape factor) is in fact twice the area of the triangle whose vertices have coordinates $(X_1, Y_1, 0) \dots (X_3, Y_3, 0)$.

S is a prime polynomial since it is of degree less or equal to 1 in any of the unknowns X_i, Y_i . P_1 is zero any time that S is null. According to theorem 2, we obtain the following factorization :

$$P_1 = S \times P_2 \quad (21)$$

P_2 is a polynomial of degree less or equal to 3 in (X, Y) and of degree less or inferior to 4 if considered as a polynomial in X_i, Y_i .

Another remark can be stated about the P_2 polynomial. Indeed, there exists another class of obvious singularities (which we will call *simple pivot singularities*). Interestingly the existence of such singularities was already mentioned in [1] and [3]. They correspond

to cases where the point with coordinates X, Y, Z (the optical center) lies on the vertical containing one of the triangle's vertices. By vertical we mean a straight line normal to the plane containing the triangle. Let us consider the rigid motion in which 2 points of the triangle are kept fixed and the last one (the one on the vertical of one vertex) moves on a circle. The tangent to this circle being vertical, the motion of this point cannot be detected on the projection on the image plane, thus it is a singularity (see figure 3).

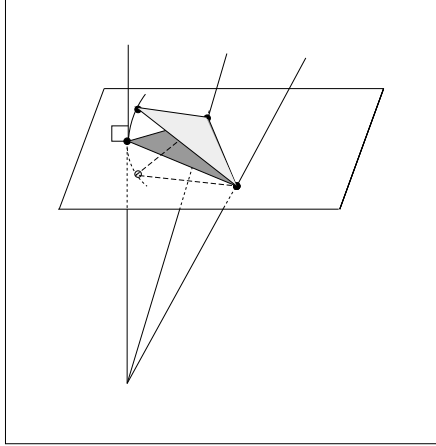


Figure 3: Simple pivot singularity

We can now say that the equation $P_2 = 0$, represents a curve in the plane (X, Y) which contains the 3 vertices of the reference triangle.

3.3.2 Result Statement

The remarkable result is that the above mentioned curve is in fact a circle. More precisely :

$$P_1 = 4 \times A^2 \times [(X - X_0)^2 + (Y - Y_0)^2 - \rho^2] \quad (22)$$

The last term is a polynomial defining an equation for the circle which contains both 3 vertices of the triangle. A is the area of this triangle.

We can state the following theorem :

Theorem 3 *The singularities that appear in the visual servoing scheme based on 3 rigidly connected points are of two natures :*

1. *The 3 points are aligned (degenerate triangle): for any position of the optical center and attitude of the coordinate frame attached to the camera, we always get a singularity.*
2. *The optical center lies on the cylinder whose axis is perpendicular to the plane containing all 3 points and which includes the 3 points: for any attitude of the coordinate frame attached to the camera the configuration is singular (see figure 4).*

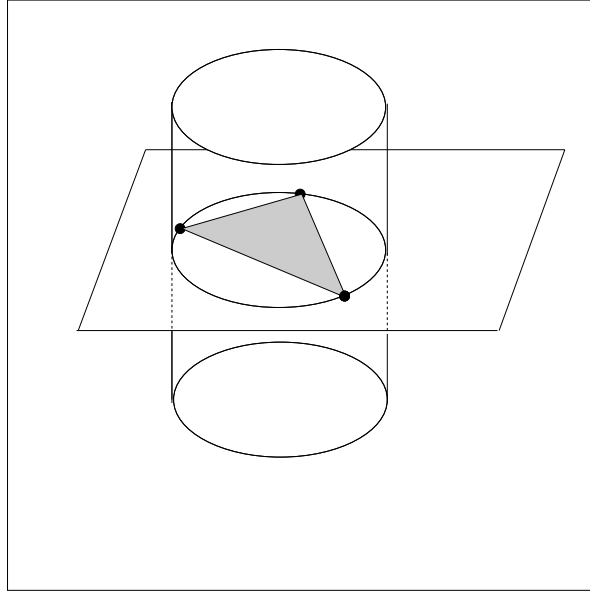


Figure 4: Cylinder of singularities

A question remains to be addressed at this point: what is the motion which would not be detected by the camera when we have a singular configuration of the second kind ?

The purely geometric proof is not obvious in the general case but if the optical center and the 3 reference points are coplanar we get the figure shown in [4]. The description given by these authors is: “if the CP (center of projection or optical center) and any number of control points lie on the same circle, then the angle between any pair of control points and the CP will be independent of the location of the CP (and hence the location of the CP cannot be determined).” This is the consequence of an elementary geometry result. In term of undetected motions, let us consider the motion whose rotation axis is the vertical containing the point which is diametrically symmetrical to the optical center, with a null sliding velocity. Since the figure contains 3 right angles, all the points move on the rays drawn from the optical center: thus this motion cannot be detected (see figure 5).

4 Discussion

4.1 The Local Determination problem (LDP)

The terms of this problem (LDP problem) are: given a triangle with sides of known lengths, recover the location of the center of projection (and orientation of the image plane if desired) from the locations of the projections of the 3 points in an image. This problem is equivalent to the “perspective-3-point” problem (P3P) as stated in [4] where it is showed that the LDP problem can be solved in 3 steps :

1. Solve the related P3P problem.

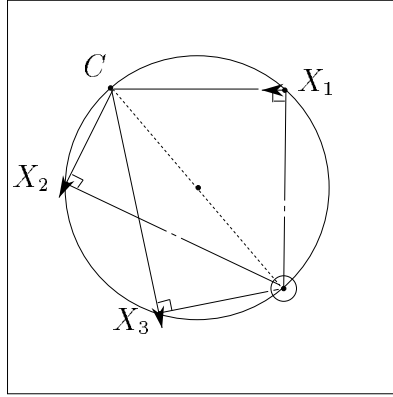


Figure 5: Undetected motion

2. Locate the center of projection with respect to the 3-D reference frame in which the control points were specified.
3. Compute the orientation of the image plane with respect to the reference frame.

In the P3P problem, the unknowns are the lengths of the line segments (“legs”) joining the center of projection to the vertices of the given triangle. The data are measures of all angles to every pair of vertices from the center of projection. As show in [4] or [9], there are in general 4 solutions to this problem. The discussion on the number of *real and positive* solutions has not been done, yet. From our simulations, all cases are possible. Given 3 angles, it is possible to find from 0 to 4 different sets of values for the “legs.” Another result available is that the similar problem (P4P) for 4 points is not ambiguous (no more than one set of solutions is possible) if the 4 points are coplanar [6].

In practice, although an algorithm to find the solutions exists, it is very sensitive to noise on the input.

4.1.1 Location of the center of projection

As explained above, this is the second part of the process described in [4]. In this reference, a rather geometrical solution is given. The mathematical expressions of the solution of this problem are extremely concise. We will express the position of the Center of Projection (CP) with :

- the barycentric coordinates μ_i (with the respect to the 3 reference points) of the orthogonal projection of the CP in the plane containing the 3 reference points.
- Z the distance between the CP and the orthogonally projected point mentioned above.

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \frac{1}{16 * A^2} \left[\begin{pmatrix} -2\Delta_1 & \alpha_3 & \alpha_2 \\ \alpha_3 & -2\Delta_2 & \alpha_1 \\ \alpha_2 & \alpha_1 & -2\Delta_3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} + \begin{pmatrix} \Delta_1\alpha_1 \\ \Delta_2\alpha_2 \\ \Delta_3\alpha_3 \end{pmatrix} \right] \quad (23)$$

$$Z^2 = \frac{1}{16 * A^2} \left[\frac{1}{2} (\lambda_1 \lambda_2 \lambda_3) \begin{pmatrix} -2\Delta_1 & \alpha_3 & \alpha_2 \\ \alpha_3 & -2\Delta_2 & \alpha_1 \\ \alpha_2 & \alpha_1 & -2\Delta_3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} + (\lambda_1 \lambda_2 \lambda_3) \begin{pmatrix} \Delta_1 \alpha_1 \\ \Delta_2 \alpha_2 \\ \Delta_3 \alpha_3 \end{pmatrix} - \Delta_1 \Delta_2 \Delta_3 \right] \quad (24)$$

with the following notations :

- Δ_i are the squared distances between pairs of vertices of the reference triangle. $\Delta_1 = d(M_2, M_3)^2, \Delta_2 = d(M_3, M_1)^2, \Delta_3 = d(M_1, M_2)^2$.

•

$$\alpha_i = \sum_{j=1}^3 \Delta_j - 2\Delta_i$$

- λ_i is the squared length of the “leg” i .
- A is the area of the reference triangle.

So, the “on the plane” coordinates of the center of projection are polynomial functions of degree one of the inputs λ_i , while the “orthogonal to the plane” coordinate, squared, is a polynomial function of degree two of the inputs.

Once again, following common intuition, we observe from these equations that when the reference triangle degenerates to a segment A becomes zero and a singularity appears, in what follows we will assume that the reference triangle is not degenerate.

The function to use to recover the location of the center of projection from the length of the “legs” is perfectly regular except when Z is zero due to the singularity of the square root function at the origin.

Given a set of lengths of the legs, which are the output of stage 1 (solving a P3P problem), we can recover the location of the center of projection up to a symmetry with respect to the plane which contains the reference triangle. The interpretation of this fact is simple : when we deal only with distances we lose the orientation information. Both symmetrical locations correspond to the same set of lengths, but in one case we would observe the reference points clockwise oriented, and in the other one counterclockwise oriented.

So, once we choose whether to operate above the plane containing the reference triangle or below, we get a one to one mapping, continuous at any order of derivation between the locations of the center of projection (\mathbb{R}^3) and a set of triplets of distances which form a subset of the set of strictly positive real numbers. The geometry of this set is easy to find from equation 24. The Matrix which appears in the right-hand side of this equation has the following properties :

- Its kernel is generated by the vector $(1, 1, 1)$.
- The two other eigenvalues are strictly negative numbers since :
 - the trace (sum of the diagonal terms) is negative.
 - their product ($48A^2$) is positive.

Since the dot product of vector $(1, 1, 1)$ and vector $(\Delta_1\alpha_1, \Delta_2\alpha_2, \Delta_3\alpha_3)$ is $16A^2$ and then strictly positive we can make a change of coordinates (from λ_i to new coordinates Λ_i), so that the right-hand side of equation 24 has the form :

$$-a^2\Lambda_1^2 - b^2\Lambda_2^2 + c^2\Lambda_3 + d$$

and the coordinate Λ_3 corresponding to an axis parallel to $(1, 1, 1)$.

Finally, since the right-hand side of equation 24 must be positive, the λ_i lies above an elliptical paraboloid whose axis is parallel to vector $(1, 1, 1)$.

4.1.2 computation of the orientation of the image frame

When we reach this step we know the coordinates of \vec{CM}_i (the “rays” from the center of projection to the reference points) in two coordinate frames :

- in a frame attached to the camera since we obtained the length of these vectors in step 1 and their directions expressed in this frame are the input of the whole process.
- in a reference frame where the coordinate of the reference 3 points are known, since we obtained the coordinates of C at the end of step 2.

To recover the orientation of the camera frame, we have to solve the following equation :

$$T_{\text{Reference Frame}} = \Theta T_{\text{Camera Frame}} \quad (25)$$

where $T_{\text{Frame f}}$ is a matrix, whose columns are the coordinates of the vectors CM_i expressed in Frame f. Θ is the rotation matrix which defines the orientation of the camera. Clearly, solving for Θ is a function of the coordinates CM_i expressed in both frames which is regular at any order of derivation, provided that the vectors CM_i are independent. Since we are solving for a rotation matrix, we can release this constraint and request only 2 among the 3 vectors to be independent, in this case we would take as third vector the vector product of the two independent vectors. The only singular case happens when the rank of the matrices T is one. In this case all three projected points in the image plane would be identical. This case is impossible since we assumed that the reference triangle was not degenerate.

So the last step : recovering the orientation only, does not add any singularity.

4.2 The “differential version” of the Local Determination problem

The process of recovering the location of the center of projection and orientation of the camera frame was decomposed in 3 steps. From this decomposition, it is easy to check that the singularities of the full process are the same as the singularities of step 1 (the P3P problem), except maybe when the CP lies in the plane containing the reference triangle. In this last case, step 2 adds singularities so it is impossible to conclude.

For the moment, we can say that the function Φ which associates the cosines of the angles between “rays” ($\widehat{M_iCM_j}$) to the location of the center of projection must be singular on the above defined “cylinder of singularities”. However, we can not decipher at this point whether the plane or parts of it correspond to singularities of this simplified problem.

The Jacobian matrix of the function Φ is very easy to obtain. Let us call ψ the function which gives the cosine of the angle between the vectors $M\vec{M}_i$ and $M\vec{M}_j$ (where M_i and M_j are two points), the value of ψ is

$$\psi(M) = \frac{M\vec{M}_i \cdot M\vec{M}_j}{\|M\vec{M}_i\| \|M\vec{M}_j\|}$$

The gradient of ψ is easy to obtain, we get :

$$\text{grad}_M \psi = \frac{1}{\|M\vec{M}_i\| \|M\vec{M}_j\|} \left[M\vec{M}_i \cdot M_i \vec{M}_j \frac{M\vec{M}_i}{\|M\vec{M}_i\|^2} + M\vec{M}_j \cdot M_j \vec{M}_i \frac{M\vec{M}_j}{\|M\vec{M}_j\|^2} \right] \quad (26)$$

If we define the function Φ as :

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{\Phi} & \mathbb{R}^3 \\ C & \longrightarrow & \begin{pmatrix} \cos(M_2 \widehat{C} M_3) \\ \cos(M_1 \widehat{C} M_3) \\ \cos(M_1 \widehat{C} M_2) \end{pmatrix} \end{array} \quad (27)$$

The determinant of the Jacobian matrix of function Φ is :

$$\begin{aligned} \det(\text{grad}(\Phi)_C) &= \frac{1}{\|C\vec{M}_1\|^2 \|C\vec{M}_2\|^2 \|C\vec{M}_3\|^2} \\ &\times \left| \frac{C\vec{M}_1}{\|C\vec{M}_1\|^2}, \frac{C\vec{M}_2}{\|C\vec{M}_2\|^2}, \frac{C\vec{M}_3}{\|C\vec{M}_3\|^2} \right| \\ &\times \begin{vmatrix} 0 & C\vec{M}_2 \cdot M_2 \vec{M}_3 & C\vec{M}_3 \cdot M_3 \vec{M}_2 \\ C\vec{M}_1 \cdot M_1 \vec{M}_3 & 0 & C\vec{M}_3 \cdot M_3 \vec{M}_1 \\ C\vec{M}_1 \cdot M_1 \vec{M}_2 & C\vec{M}_2 \cdot M_2 \vec{M}_1 & 0 \end{vmatrix} \end{aligned} \quad (28)$$

We used the notation $|V_1, V_2, V_3|$ for the determinant of the vectors V_1, V_2, V_3 and $|M|$ for the determinant of a matrix M .

From equation 29, singularities for the function Φ occur in two cases :

1. When the determinant of the 3 vectors $C\vec{M}_i$ is zero which is equivalent to the condition that M belongs to the plane which contains the reference triangle (the “reference plane”. If the reference triangle is degenerate, this determinant is always zero.
2. we recognize the last term : it is the determinant of the “characteristic matrix” that we encountered previously, its value is zero if C belongs to the “cylinder of singularities”

We can state the following theorem :

Theorem 4 *The singularities that appear in the computation of the location of the center of projection (CP) , given the cosines of the angle between all 3 pairs of “rays” which connect the center of projection to one among 3 reference points are :*

- any location of the CP. if the 3 points are aligned (degenerate triangle).

- If the reference triangle is not degenerate there are two cases of singularity :
 1. the CP belongs to the plane which contains the 3 reference points (“reference plane”).
 2. the CP belongs to the circular cylinder whose axis is perpendicular to the reference plane and which contains all 3 points.

4.2.1 The “paradox” of the reference plane singularity

It may seem strange at first sight, that the reference plane corresponds to the singularities for the restricted local determination problem, though it does not for the complete problem.

To give an intuitive understanding of this fact, let us consider a location of the center of projection which lies in the “reference plane” without lying on the “cylinder of singularities”. The kernel of the Jacobian matrix of Φ is of dimension 1 : it corresponds to motions of the CP which are orthogonal to the reference plane. If we move “vertically” from a location on the “reference plane”, we would not observe a corresponding infinitely small change in the cosines.

Let us now switch to the complete problem. In the case where the center of projection lies on the “reference plane”, the 3 projected points are aligned in the image. We can suppose that this line is the x axis. The optic flow equations have the form :

$$\begin{pmatrix} \dot{x}_i \\ \dot{y}_i \end{pmatrix} = - \begin{pmatrix} \frac{1}{Z_i} & 0 & -\frac{x_i}{Z_i} & 0 & 1 + x_i^2 & 0 \\ 0 & \frac{1}{Z_i} & 0 & -1 & 0 & x_i \end{pmatrix} \cdot T \quad (29)$$

We are now interested in finding infinitely small motions, which would produce null \dot{x}_i and \dot{y}_i . Since this implies also null derivate of the cosines of the angles between “rays”, it follows from a previous paragraph that all translation components must be null except maybe the component orthogonal to the “reference plane”. This corresponds to null V_x and V_z with, maybe, a non null V_y . From the equation for \dot{x}_i , Ω_y must be null.

Now, let us consider the equation in \dot{y}_i , unless the Z_i are affine functions of the x_i (which does not occur since the reference triangle is not degenerate), there is no corresponding choice of V_y , Ω_x and Ω_z (which are not all null) to obtain a null \dot{y}_i . In even more schematic words, since the reference points have different depths, a translation of the CP orthogonal to the “reference plane”, even when combined with an arbitrary rotation of the camera can not produce all null \dot{y}_i .

Finally, we can conclude that, in general, for a center of projection lying on the “reference plane”, we can not find a motion (a set of V_x , V_y . . . Ω_z not all null) which produces null \dot{x}_i and \dot{y}_i . In general the “reference plane” is not singular for the complete problem.

4.3 Relation between the singularities and the local determination problem

Following common robotics experience, it seemed interesting to find a relation between the position of *equivalent configurations* (positions in 6 dimensional space that give the same set of image features) and the locations of singular configurations.

Let us consider the case of the simple two-joint planar mechanism for example. And consider the *geometric model* which gives the cartesian coordinates of the extremity of the last bar O_2 as a function of the two angles (q_1, q_2) (see figure 6).

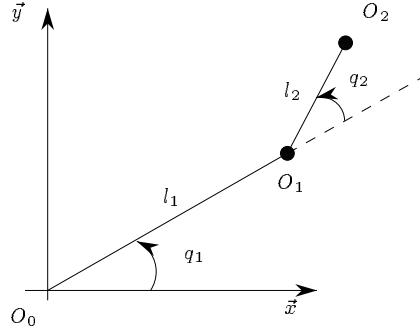


Figure 6: 2 Joints Planar Mechanism

For this mechanism, *equivalent configurations* are sets of angles q_1, q_2 which all give the same coordinates for the extremity of the last bar. It is easy to show that in general, if two sets of joints coordinates are equivalent then they have opposite q_2 coordinates. The only peculiar case is when the lengths of the two arms are equal. In this case, all angles of the form q_1, π correspond to the origin O_0 . The singular configurations are those for which $q_2 = 0$ or $q_2 = \pi$.

So, we can state a very interesting result for this mechanism (true for any non zero lengths of the two bars). If there exists a path joining two equivalent configurations then this path necessarily encounters a singular configuration. This peculiarity is true for other classical mechanisms. The practical consequence is that though there exist equivalent configurations, the situation is not so dreadful since two equivalent configurations are always separated by a “singularity barrier.”

Unfortunately that is not the case for our problem. Let us consider the example of 4 equivalent configurations given in [4]. This example correspond to an equilateral triangle. The first configuration corresponds to a tetrahedron with equal “legs”, the center of projection is *inside* the singularities cylinder (see figure 7).

There are 3 other equivalent configurations. They can all be obtained after applying a subsequent rotation of angle $\frac{2\pi}{3}$ or $-\frac{2\pi}{3}$ on the center of projection once one of these configurations is found. For these last 3 configurations, the center of projection is *outside* the singularities cylinder(see figure 8).

We can thus define a path joining any couple of these 3 configurations which remains *outside* the singularities cylinder and thus does not encounter any “singularity barrier”.

At this point there is only one clear relation that we have seen between the equivalent solutions to the P3P problem and the singularities problem. A singular configuration corresponds to a case where there is a double root for the fourth degree polynomial that we encounter in the algorithm that solves the P3P problem.

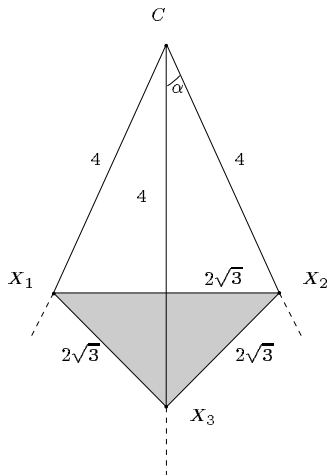


Figure 7: A configuration for the P3P problem

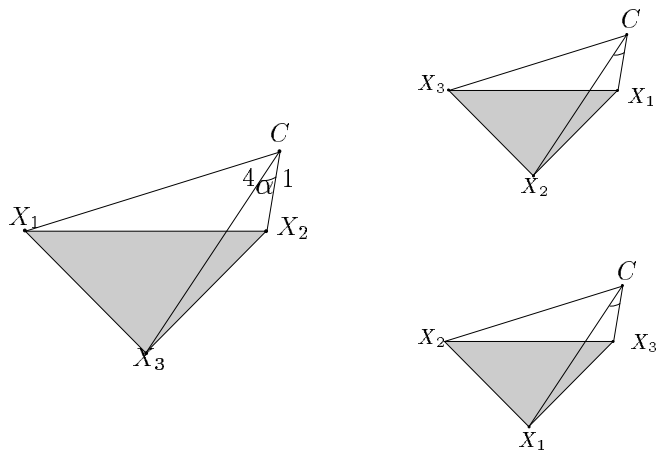


Figure 8: 3 Others equivalent configuration for the P3P problem

5 Conclusion

The 3 points perspective singularities problem that we solved in this article is probably easier than similar ones based on the projection of complex algebraic curves. The interesting peculiarity is that the rigid motion constraint is expressed naturally in the representation of the problem itself. This constraint is simply formulated by the choice of 3 distances in 3D space. In other similar problems, it is much more difficult to express that 2 images correspond to the same rigid body in a concise algebraic equation.

However we believe that this approach contain two ideas that can be generalized. First, it seems very likely that similar problems based on algebraic curves can be solved with similar algebraic considerations in mind. Second, the most fundamental finding is that the singular nature of the situation of the camera coordinate frame depends only on the position of the optical center. This idea seems obvious if we think in terms of the image sphere instead of the image plane. Our intuition is that a view of an object can always be represented by two quantities. First a representation of a rotation on the sphere image (that is to say a representation of $SO3$ and then an intrinsic representation of the shape and size of the “visual cone” under which the object is perceived. In our peculiar problem the “visual cone” is the tetrahedron whose vertices are the 3 reference points and the center of projection. In this case, a rotation on the image sphere does not change the singular nature of a configuration. This is probably general since the fact of being a projective singularity is an intrinsic characteristic of a problem which does not depend on subsequent regular remapping of the image sphere.

We fully realize that we have addressed only one of the requirements which must be fulfilled so that a set of visual features are acceptable to perform a visual servoing task. Other conditions are of importance for control. Several were mentioned in [3]. For example, norm of the Jacobian matrix should not be either too large for control’s sensitivity to noise considerations or too small since the camera would not then be able to accurately sense the motion. So, similar results that the one we obtained for the determinant of the interaction matrix must be found for other quantities. Examples of such quantities are the norm of the Jacobian matrix or its inverse, the condition of this matrix, a measure of the variation of the matrix with the displacement of the optical center. However, we expect such findings only for quantities which have an intrinsic nature or *invariance properties*. According to the *Weyl’s thesis* often referred to in [8] those invariant quantities are the ones that have natural physical interpretation or geometrical “meaning”.

References

- [1] F. Chaumette: *La relation vision-commande: théorie et applications à des tâches robotiques*, Ph-D Thesis, University of Rennes I, France, July 1990.
- [2] B. Espiau, F. Chaumette, P. Rives: *A new approach to visual servoing in robotics*, Accepted for publication in IEEE Transactions on Robotics and Automation.
- [3] J. T. Feddema, C. S. G. Lee and O. R. Mitchell: *Weighted selection of image features for resolved rate visual feedback control*, IEEE Transactions on Robotics and Automation, Vol. 7, n.1, pp. 31-47, February 1991.

- [4] M. A. Fischler and R. C. Bolles : *Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography*, Communications ACM, Vol. 44, pp. 381-395, 1981.
- [5] A. L. Gilbert, K.G. Giles, G. M. Flachs, R. B. Rogers, Y. H. U : *A Real-Time Video Tracking System*, IEEE Transaction on PAMI, Vol. 2, n.1, pp. 47-56, January 1980.
- [6] R. Horaud, B. Conio, O. Lebouleux : *An analytic solution for the perspective 4-Point Problem*. Computer Vision, Graphics, and Image Processing **47**, pp. 33-44, 1989.
- [7] M. Kabuka, E. McVey, P. Shironoshita : *An Adaptive Approach to Video Tracking*, IEEE Journal of Robotics and Automation, Vol. 4, n. 2, pp. 228-236, April 1988.
- [8] K. Kanatani : *Group-theoretical methods in image understanding*. Springer-Verlag, Berlin Heidelberg, Germany, 1990.
- [9] S. Linnainmaa, D. Harwood, L. S. Davis : *Pose determination of a three-dimensional object using triangle pairs*. IEEE Transaction on PAMI, Vol. 10, n.5, pp. 634-647, September 1988.
- [10] R. P. Paul, *Robot Manipulators : Mathematics, Programming, and Control*. Cambridge, Ma : MIT Press, 1981.
- [11] C. Samson, B. Espiau, M. Le Borgne: *Robot Control: the Task Function Approach*, Oxford University Press, 1991.
- [12] A. C. Sanderson, L. E. Weiss : *Image Based Visual Servo Control Using Relational Graph Error Signal*, Proc. of the Int. Conf. on Cybernetics and Society, Cambridge, MA, IEEE SMC, pp. 1074-1077, October 1980.
- [13] L. E. Weiss, A. C. Sanderson: *Dynamic Sensor-Based Control of Robots with Visual Feedback*, IEEE Journal of Robotics and Automation, Vol. RA-3, n. 5, Oct. 1987.