Rigid Transformation Estimation in Point Registration Procedures: Study of the Use of Mahalanobis Distances

Manuel Yguel

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Estimation de la transformation rigide dans les procédures d'alignement de points: étude du cas de l'utilisation de distances de Mahalanobis

Manuel Yguel

N° 6684

Octobre 2008

Thèmes COG et NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6684--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
GRENOBLE - RHÔNE-ALPES

Estimation de la transformation rigide dans les procédures d'alignement de points: étude du cas de l'utilisation de distances de Mahalanobis

Manuel Yguel*

Thèmes COG et NUM — Systèmes cognitifs et Systèmes numériques
Équipe-Projet e-motion

Rapport de recherche n° 6684 — Octobre 2008 — 21 pages

Résumé : Ce document s'intéresse au problème d'estimation de la transformation rigide qui minimise la somme des carrés des distances d'appariement d'un ensemble de couples de points de dimension n lorsque sont utilisées des distances de Mahalanobis. Ce problème est énnoncé comme un problème de minimization sous contraintes et la technique des multiplicateurs de Lagrange est utilisée. La solution de ce problème est donnée sous la forme d'un système d'équations polynomiales. Parmi les solutions de ce système se trouve la solution du problème d'estimation. Les cas 2-D et 3-D sont étudiés explicitement et des systèmes d'équations simplifiés sont présentés pour ces cas précis ainsi que des résultats d'optimalité concernant le choix des multiplicateurs de Lagrange.

Mots-clés : alignement de points, transformation rigide, distance de Mahalanobis, ICP, ECMPR, résolution de systèmes polynomiaux

* INPG PhD applicant, e-motion team, grant from UJF

Centre de recherche INRIA Grenoble – Rhône-Alpes
655, avenue de l'Europe, 38334 Montbonnot Saint Ismier
Téléphone : +33 4 76 61 52 00 — Télécopie +33 4 76 61 52 52

Rigid Transformation Estimation in Point Registration Procedures: Study of the Use of Mahalanobis Distances

Abstract: This document explores the problem of the estimation of the rigid body transformation that minimizes the sum of squares between matched points of dimension n where Mahalanobis distances are used. This problem is stated as a constrained minimization problem and the Lagrange multipliers framework is used. The solution is given as a system of polynomial equations. Among the solution of that system is the solution of the estimation problem. The 2-D and the 3-D cases are explicitely studied and simplified equations are given for those cases. In addition results are given concerning the optimal choice of the Lagrange multiplier.

Key-words: point registration, rigid body transformation, Mahalanobis distance, ICP, ECMPR, polynomial systems

1 Estimating the optimal rigid body transformation using Mahalanobis distances

1.1 Problem setting

Assume that there exist two corresponded point sets $\{\mathbf{p}_i\}$ and $\{\mathbf{q}_i\}$ along with a set of Mahalanobis $n \times n$ matrices $\mathbf{\Pi}_i$, $i = 1 \dots N$, such that they are related by :

$$\mathbf{q}_i \quad \longleftarrow \quad \mathbf{R}\mathbf{p}_i + \mathbf{t} + \mathbf{v}_i \tag{1}$$

where $\mathbf{R}$ is a standard $n \times n$ rotation matrix, $\mathbf{t}$ is a n-D translation vector and $\mathbf{v}_i$ a n-D noise vector.

The objective is to find the optimal transformation $[\mathbf{R}^*, \mathbf{t}^*]$ that maps the set $\{\mathbf{p}_i\}$ onto $\{\mathbf{q}_i\}$ which minimizes the *Mahalanobis least squares error criterion* given by :

$$\Sigma^2(\mathbf{R}, \mathbf{t}) \quad \triangleq \quad \sum_{i=1}^{N} \mathbf{v}_i^T \mathbf{\Pi}_i \mathbf{v}_i = \sum_{i=1}^{N} [\mathbf{q}_i - (\mathbf{R}\mathbf{p}_i + \mathbf{t})]^T \, \mathbf{\Pi}_i \, [\mathbf{q}_i - (\mathbf{R}\mathbf{p}_i + \mathbf{t})] \tag{2}$$

For a point $\mathbf{R}\mathbf{p}_i + \mathbf{t}$, the projection matrix, $\mathbf{\Pi}_i$ authorizes, for instance, the point $\mathbf{q}_i$ to be on a line or a plane since the Mahalanobis projection can assign a small or zero weight to a displacement along such a line or plane. Using Mahalanobis distances is more general, it allows to give n different wheights for the displacements in the n orthogonal directions of any reference system. It is possible to penalize and authorize displacements locally knowing the local geometry of the objects. Therefore the point registration process is more robust since two different points of the same linear subspace, for instance, can be associated. It is also more realistic : the Euclidean distance model assume that it is possible to model the world with points. However this kind of model is true only if an infinite precision and memory is available to encode the world representation. Therefore in the Euclidean approache an incompleteness problem arises du to the finite representation : the points of the second cloud never correspond to the points of the first cloud.

To ensure that $\mathbf{R}$ is a rotation matrix it is equivalent that two conditions are fulfilled :

$$\mathbf{R}^T\mathbf{R} \quad = \quad \mathbf{I} \tag{3}$$

$$\det(\mathbf{R}) \quad = \quad 1 \tag{4}$$

The first equation (eq. 3) states that the rotation matrix is orthogonal and the second one (eq. 4) implies that there is a conservation of the angles between figures along the transformation.
This is now possible to cast the problem into a constrained minimization framework which Lagrangian is :

$$\mathscr{L}(\mathbf{R}, \mathbf{t}) \quad = \quad \Sigma^2(\mathbf{R}, \mathbf{t}) + \mathrm{Tr}\ \left(\mathbf{L}[\mathbf{R}^T\mathbf{R} - \mathbf{I}]\right) + \lambda_d(det(\mathbf{R}) - 1) \tag{5}$$

where $\mathbf{L}$ is a matrix of Lagrange multipliers.

Now the objective is to find the zero of the derivatives of $\mathscr{L}(\mathbf{R}, \mathbf{t})$ with respect to $\mathbf{t}$ and $\mathbf{R}$.

1.2 $\mathbf{t}^*$ as a function of $\mathbf{R}^*$

As for the Euclidean case, there are no constraints on $\mathbf{t}$. Therefore it is possible to get rid of $\mathbf{t}$ by substituting an expression of $\mathbf{t}^*$ as a function of $\mathbf{R}^*$.

$$\begin{aligned} \frac{\partial\ \mathscr{L}(\mathbf{R}, \mathbf{t})}{\partial\ \mathbf{t}} &= \frac{\partial\ \Sigma^2(\mathbf{R}, \mathbf{t})}{\partial\ \mathbf{t}} \\ &= -2 \sum_j \mathbf{\Pi}_j \left[\mathbf{q}_j - (\mathbf{R}\mathbf{p}_j + \mathbf{t}) \right] \end{aligned}$$

Thus, setting $\frac{\partial\ \mathscr{L}(\mathbf{R}, \mathbf{t})}{\partial\ \mathbf{t}}$ to zero leads to the following equation :

$$\sum_j \mathbf{\Pi}_j \left[\mathbf{q}_j - \mathbf{R}\mathbf{p}_j \right] - \left[\sum_j \mathbf{\Pi}_j \right] \mathbf{t} = \mathbf{0}$$

and therefore $\mathbf{t}^*$ can be expressed as a function of $\mathbf{R}^*$:

$$\mathbf{t}^* = \left[\sum_j \mathbf{\Pi}_j \right]^{-1} \sum_j \mathbf{\Pi}_j \left[\mathbf{q}_j - \mathbf{R}^* \mathbf{p}_j \right] \tag{6}$$

In the following we note :

$$\mathbf{K} = \left[\sum_j \mathbf{\Pi}_j \right]^{-1} \tag{7}$$

$$\mathbf{l} = \sum_j \mathbf{K}\mathbf{\Pi}_j \mathbf{q}_j \tag{8}$$

We remark that as each $\mathbf{\Pi}_i$ is symmetric, $\sum_j \mathbf{\Pi}_j$ is also symmetric and necessarily, $\mathbf{K}$ is also symmetric. For the same reasons, $\mathbf{K}$ is also semi-definite positive or definite positive wether the $\mathbf{\Pi}_i$'s are semi-definite or definite.

With these notations :

$$\mathbf{t}^* = \mathbf{l} - \mathbf{K} \sum_j \mathbf{\Pi}_j \mathbf{R}^* \mathbf{p}_j . \tag{9}$$

1.3 Expansion of Σ^2

We drop the star in the notations of $\mathbf{t}^*$ and $\mathbf{R}^*$ until the end of the section for clarity. But as we have for a given rotation the optimal translation, we search to minimize a Σ^2 that takes the optimal translation into account.

Let us note

$$\mathbf{x}_i = \mathbf{q}_i - (\mathbf{R}\mathbf{p}_i + \mathbf{t})$$

such that $\Sigma^2 = \sum_i \mathbf{x}_i^T \mathbf{\Pi}_i \mathbf{x}_i$. And let us replace $\mathbf{t}$ by its expression as a function of $\mathbf{R}$ (eq. 9) in $\mathbf{x}_i$:

$$\mathbf{x}_i = \mathbf{q}_i - \mathbf{R}\mathbf{p}_i - \mathbf{l} + \sum_j \mathbf{K}\mathbf{\Pi}_j \mathbf{R}\mathbf{p}_j.$$

1.3.1 Expansion of $\mathbf{x}_i^T \mathbf{\Pi}_i \mathbf{x}_i$

Let us note, temporary,

$$\mathbf{m} = \sum_j \mathbf{K}\mathbf{\Pi}_j \mathbf{R}\mathbf{p}_j \tag{10}$$

such that :

$$\mathbf{x}_i = \mathbf{q}_i - \mathbf{R}\mathbf{p}_i - \mathbf{l} + \mathbf{m}.$$

Now we can expand $\mathbf{x}_i^T \mathbf{\Pi}_i \mathbf{x}_i$. We obtain 16 summands that we can group into 10 ones thanks to the fact that $\mathbf{x}^T \mathbf{\Pi}_i \mathbf{y} = \mathbf{y}^T \mathbf{\Pi}_i \mathbf{x}$ since $\mathbf{\Pi}_i$ is symmetric.

The first group of 3 summands does not contain $\mathbf{R}$:

$$\kappa_i = \mathbf{q}_i^T \mathbf{\Pi}_i \mathbf{q}_i - 2\mathbf{q}_i^T \mathbf{\Pi}_i \mathbf{l} + \mathbf{l}^T \mathbf{\Pi}_i \mathbf{l} \tag{11}$$

Each summand in the second group of 7 contains $\mathbf{R}$ and can be rewritten using Kronecker (see 5.1) and vec operators (see 5.2) :

$$-2\mathbf{q}_i^T \mathbf{\Pi}_i \mathbf{R}\mathbf{p}_i = -2\mathbf{q}_i^T \left[\mathbf{p}_i^T \otimes \mathbf{\Pi}_i\right] \operatorname{vec}(\mathbf{R}) \tag{12}$$

$$2\mathbf{q}_i^T \mathbf{\Pi}_i \mathbf{m} = 2\mathbf{q}_i^T \mathbf{\Pi}_i \mathbf{K}\mathbf{M} \operatorname{vec}(\mathbf{R}) \tag{13}$$

$$(\mathbf{R}\mathbf{p}_i)^T \mathbf{\Pi}_i \mathbf{R}\mathbf{p}_i = \operatorname{vec}(\mathbf{R})^T \left[\mathbf{p}_i \mathbf{p}_i^T \otimes \mathbf{\Pi}_i\right] \operatorname{vec}(\mathbf{R}) \tag{14}$$

$$-2(\mathbf{R}\mathbf{p}_i)^T \mathbf{\Pi}_i \mathbf{m} = -2 \operatorname{vec}(\mathbf{R})^T \mathbf{M}^T \mathbf{K}^T \left[\mathbf{p}_i^T \otimes \mathbf{\Pi}_i\right] \operatorname{vec}(\mathbf{R}) \tag{15}$$

$$2\mathbf{l}^T \mathbf{\Pi}_i \mathbf{R}\mathbf{p}_i = 2\mathbf{l}^T \left[\mathbf{p}_i^T \otimes \mathbf{\Pi}_i\right] \operatorname{vec}(\mathbf{R}) \tag{16}$$

$$-2\mathbf{l}^T \mathbf{\Pi}_i \mathbf{m} = -2\mathbf{l}^T \mathbf{\Pi}_i \mathbf{K}\mathbf{M} \operatorname{vec}(\mathbf{R}) \tag{17}$$

$$\mathbf{m}^T \mathbf{\Pi}_i \mathbf{m} = \operatorname{vec}(\mathbf{R})^T \mathbf{M}^T \mathbf{K}^T \mathbf{\Pi}_i \mathbf{K}\mathbf{M} \operatorname{vec}(\mathbf{R}) \tag{18}$$

where

$$\mathbf{M} = \sum_j \mathbf{p}_j^T \otimes \mathbf{\Pi}_j. \tag{19}$$

All these relations are based upon the following facts, using eq. 75 then theorem 7 :

$$\mathbf{R}\mathbf{p}_i = \operatorname{vec}(\mathbf{R}\mathbf{p}_i) = (\mathbf{p}_i^T \otimes \mathbf{I}_{n\times n}) \operatorname{vec}(\mathbf{R}). \tag{20}$$

Moreover :

$$\begin{aligned} \mathbf{\Pi}_i(\mathbf{p}_i^T \otimes \mathbf{I}_{n\times n}) &= \mathbf{p}_i^T \otimes \mathbf{\Pi}_i & (21)\\ \left[\mathbf{p}_i \otimes \mathbf{I}_{n\times n}\right]\left[\mathbf{p}_i^T \otimes \mathbf{\Pi}_i\right] &= \mathbf{p}_i\mathbf{p}_i^T \otimes \mathbf{\Pi}_i & (22)\end{aligned}$$

such that using in addition theorem (5), all the preceding equations are easily derived.

Using theorem (5) and theorem (6), it is possible to rewrite eq. (12). Since :

$$\begin{aligned} \mathbf{q}_i^T\left[\mathbf{p}_i^T \otimes \mathbf{\Pi}_i\right] &= \left(\left[\mathbf{p}_i \otimes \mathbf{\Pi}_i\right]\mathbf{q}_i\right)^T \\ &= \left[\ \mathrm{vec}\left(\mathbf{\Pi}_i\mathbf{q}_i\mathbf{p}_i^T\right)\right]^T & (23)\end{aligned}$$

eq. (12) is :

$$-2\mathbf{q}_i^T\mathbf{\Pi}_i\mathbf{R}\mathbf{p}_i = -2\left[\ \mathrm{vec}\left(\mathbf{\Pi}_i\mathbf{q}_i\mathbf{p}_i^T\right)\right]^T\ \mathrm{vec}\left(\mathbf{R}\right)\ . \quad (24)$$

1.3.2 Sum over i

To obtain Σ^2, let sum over i the preceeding expressions :

- summing over i eq. (13), $\sum_i \mathbf{q}_i^T\mathbf{\Pi}_i\mathbf{K}$ appears which is $\left[\sum_i \mathbf{K}\mathbf{\Pi}_i\mathbf{q}_i\right]^T$ and one recognizes $\mathbf{l}^T$ (eq. 8) such that :

$$\sum_i 2\mathbf{q}_i^T\mathbf{\Pi}_i\mathbf{K}\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) = 2\mathbf{l}^T\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) \quad (25)$$

- summing over i eq. (15), one recognizes $\mathbf{M}$ such that :

$$\sum_i -2\ \mathrm{vec}\left(\mathbf{R}\right)^T\mathbf{M}^T\mathbf{K}^T\left[\mathbf{p}_i \otimes \mathbf{\Pi}_i\right]\ \mathrm{vec}\left(\mathbf{R}\right) = -2\ \mathrm{vec}\left(\mathbf{R}\right)^T\mathbf{M}^T\mathbf{K}\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) \quad (26)$$

- summing over i eq. (16), one also recognizes $\mathbf{M}$ such that :

$$\sum_i 2\mathbf{l}^T\left[\mathbf{p}_i \otimes \mathbf{\Pi}_i\right]\ \mathrm{vec}\left(\mathbf{R}\right) = 2\mathbf{l}^T\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) \quad (27)$$

- summing over i eq. (17), one recognizes $\mathbf{K}^{-1} = \sum_i \mathbf{\Pi}_i$ such that :

$$\sum_i -2\mathbf{l}^T\mathbf{\Pi}_i\mathbf{K}\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) = -2\mathbf{l}^T\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) \quad (28)$$

- same equality in eq. (18) gives :

$$\sum_i\ \mathrm{vec}\left(\mathbf{R}\right)^T\mathbf{M}^T\mathbf{K}^T\mathbf{\Pi}_i\mathbf{K}\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) = \mathrm{vec}\left(\mathbf{R}\right)^T\mathbf{M}^T\mathbf{K}\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right) \quad (29)$$

Eq. 27 and eq. 28 make zero ; eq. 29 and eq. 26 makes $-\ \mathrm{vec}\left(\mathbf{R}\right)^T\mathbf{M}^T\mathbf{K}\mathbf{M}\ \mathrm{vec}\left(\mathbf{R}\right)$ such that :

Theorem 1 **(Σ^2 as a function of vec$(\mathbf{R})$)** *It is possible to derive Σ^2 for a given rotation and the associated optimal translation :*

$$\Sigma^2 = vec(\mathbf{R})^T \left[\mathbf{N} - \mathbf{M}^T\mathbf{K}\mathbf{M}\right] vec(\mathbf{R}) + 2\left[\mathbf{l}^T\mathbf{M} - \mathbf{o}^T\right] vec(\mathbf{R}) + \kappa$$

where

$$\mathbf{K} = \left[\sum_i \mathbf{\Pi}_i\right]^{-1}$$

$$\mathbf{l} = \mathbf{K}\sum_i \mathbf{\Pi}_i \mathbf{q}_i$$

$$\mathbf{M} = \sum_i \mathbf{p}_i^T \otimes \mathbf{\Pi}_i$$

$$\mathbf{N} = \sum_i \mathbf{p}_i\mathbf{p}_i^T \otimes \mathbf{\Pi}_i$$

$$\mathbf{o} = vec\left(\sum_i \mathbf{\Pi}_i\mathbf{q}_i\mathbf{p}_i^T\right)$$

and κ is a constant[a] *that does not depend on* $\mathbf{R}$.

[a] $\kappa = \sum_i \kappa_i$ as defined in eq. (11)

With these notations the translation is :

$$\mathbf{t}^* = \mathbf{l} - \mathbf{K}\mathbf{M}\ \mathrm{vec}(\mathbf{R}^*)\ . \tag{30}$$

Let us note :

$$\mathbf{A} \triangleq \left[\mathbf{N} - \mathbf{M}^T\mathbf{K}\mathbf{M}\right] \tag{31}$$

$$\mathbf{b} \triangleq \left[\mathbf{o} - \mathbf{M}^T\mathbf{l}\right] \tag{32}$$

using these matrices the constrained minimization problem is stated as :

$$\begin{cases} \mathbf{R}^* = \underset{\mathrm{vec}(\mathbf{R})}{\arg\min}\ \mathrm{vec}(\mathbf{R})^T\,\mathbf{A}\ \mathrm{vec}(\mathbf{R}) - 2\mathbf{b}^T\ \mathrm{vec}(\mathbf{R})\ , \\ \\ \text{with constraints :} \\ \mathbf{R}^{*T}\mathbf{R}^* = \mathbf{I}_{n\times n}\ , \\ \det \mathbf{R}^* = 1\ . \end{cases} \tag{33}$$

One can notice several wothwhile facts :
- $\mathbf{A}$ is symmetric,
- $\mathbf{A}$ is semi-definite positive.
- the gradient of Σ^2 with respect to vec$(\mathbf{R})$ is :

$$\frac{\partial\ \Sigma^2(\mathrm{vec}(\mathbf{R}))}{\partial\ \mathrm{vec}(\mathbf{R})} = 2\ \mathrm{vec}(\mathbf{R})^T\,\mathbf{A} - 2\mathbf{b}^T \tag{34}$$

Proof:

The first property comes from direct calculus : $\mathbf{N}$ is symmetric by definition and as $\mathbf{K}$ is symmetric, $\mathbf{M}^T\mathbf{K}\mathbf{M}$ is also symmetric and so is $\mathbf{A}$.

The second property comes from the fact that $\Sigma^2(\text{ vec}\,(\mathbf{R})) \geq 0$.
If $\mathbf{A}$ were not semi-definite positive, one could find an eigenvector $\mathbf{u}$ of $\mathbf{A}$ with a negative eigenvalue $-\mu$. Theorem 1 states that

$$\Sigma^2(\mathbf{u}) = -\mu\|\mathbf{u}\|^2 - 2\mathbf{b}^T\mathbf{u} + \kappa \tag{35}$$

which is quadratic with respect to the coefficients of $\mathbf{u}$. Thus it would exist a vector $\alpha\mathbf{u}$ colinear with $\mathbf{u}$ with sufficiently great length that would made $\Sigma^2(\alpha\mathbf{u}) < 0$, since at the limit only the coefficient of $\|\mathbf{u}\|^2$ matters. And so we have a contradiction.

The third property comes from differential calculus and the fact that $\mathbf{A}$ is symmetric.

■

Theorem 2 (Equivalence with a least-squares formulation) *The minimization of Σ^2 as a function of vec$\,(\mathbf{R})$ has an equivalent least-squares formulation :*

$$\left\{\begin{array}{lcl} \mathbf{R}^* & = & \underset{vec(\mathbf{R})}{\arg\min}\|\mathbf{F}\ vec\,(\mathbf{R}) - \mathbf{u}\|^2 \ , \\ \\ \multicolumn{3}{l}{\textit{with constraints :}} \\ \mathbf{R}^{*T}\mathbf{R}^* & = & \mathbf{I}_{n\times n} \ , \\ \det\mathbf{R}^* & = & 1 \ . \end{array}\right.$$

with

$$\mathbf{A} = \mathbf{F}^T\mathbf{F}$$
$$\mathbf{b} = \mathbf{F}^T\mathbf{u}$$

Proof:

Let us rewrite :

$$\begin{aligned} \|\mathbf{F}\ \text{vec}\,(\mathbf{R}) - \mathbf{u}\|^2 &= [\mathbf{F}\ \text{vec}\,(\mathbf{R}) - \mathbf{u}]^T[\mathbf{F}\ \text{vec}\,(\mathbf{R}) - \mathbf{u}] \\ &= \text{vec}\,(\mathbf{R})^T\,\mathbf{F}^T\mathbf{F}\ \text{vec}\,(\mathbf{R}) - 2\mathbf{u}^T\mathbf{F}\ \text{vec}\,(\mathbf{R}) + \mathbf{u}^T\mathbf{u} \end{aligned}$$

Therefore minimizing $\|\mathbf{F}\ \mathrm{vec}\,(\mathbf{R}) - \mathbf{u}\|^2$ with respect to $\mathrm{vec}\,(\mathbf{R})$ is equivalent to minimizing :

$$\mathrm{vec}\,(\mathbf{R})^T\,\mathbf{F}^T\mathbf{F}\ \mathrm{vec}\,(\mathbf{R}) - 2\mathbf{u}^T\mathbf{F}\ \mathrm{vec}\,(\mathbf{R})\ .$$

Since $\mathbf{A}$ is symmetric semi-definite positive, it can be factorized as $\mathbf{F}^T\mathbf{F}$, therefore a direct identification with eq. (33) gives :

$$\mathbf{A} = \mathbf{F}^T\mathbf{F}\ ,$$

$$\mathbf{b} = \mathbf{F}^T\mathbf{u}\ .$$

■

This makes the link with the work in Viklands (2006) where the general least-squares problem, presented above, was approached using a numerical method close to that described in the following.

2 3-D Minimization

For any rotation matrix in 3-D there exists two unit quaterions. Let

$$\boldsymbol{\xi} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} + w$$

a quaternion such that $x^2 + y^2 + z^2 + w^2 = 1$ then the equivalent rotation matrix is

$$\mathbf{R}(\boldsymbol{\xi}) = \begin{bmatrix} 1-2y^2-2z^2 & 2xy-2zw & 2xz+2yw \\ 2xy+2zw & 1-2x^2-2z^2 & 2yz-2xw \\ 2xz-2yw & 2yz+2xw & 1-2x^2-2y^2 \end{bmatrix} \tag{36}$$

This is now possible to cast the problem into a constrained minimization framework which Lagrangian is :

$$\mathscr{L}(\boldsymbol{\xi}, \mathbf{t}) \quad = \quad \Sigma^2(\mathbf{R}(\boldsymbol{\xi}), \mathbf{t}) + \lambda(\boldsymbol{\xi}^T\boldsymbol{\xi} - 1) \tag{37}$$

The problem is far more simpler in terms of the number of equations, since only one Lagrange multiplier is used.

2.1 Equivalence with a polynomial system

Knowing the gradient of Σ^2 with respect to $\mathrm{vec}\,(\mathbf{R})$ (eq. 34), we use the chain rule and the Jacobian of the change of variable from $\boldsymbol{\xi}$ to $\mathrm{vec}\,(\mathbf{R})$:

$$\mathbf{D}_{\text{vec}(\mathbf{R})(\boldsymbol{\xi})} = 2 \begin{bmatrix} 0 & -2y & -2z & 0 \\ y & x & w & z \\ z & -w & x & -y \\ y & x & -w & -z \\ -2x & 0 & -2z & 0 \\ w & z & y & x \\ z & w & x & y \\ -w & z & y & -x \\ -2x & -2y & 0 & 0 \end{bmatrix} \triangleq 2\mathbf{J} , \tag{38}$$

to find the gradient of Σ^2 with respect to $\boldsymbol{\xi}$:

$$\begin{aligned} \frac{\partial\, \Sigma^2(\boldsymbol{\xi})}{\partial\, \xi} &= \frac{\partial\, \Sigma^2(\,\text{vec}\,(\mathbf{R}))}{\partial\,\text{vec}\,(\mathbf{R})} \mathbf{D}_{\text{vec}(\mathbf{R})(\boldsymbol{\xi})} \\ &= 4\left[\,\text{vec}\,(\mathbf{R}(\boldsymbol{\xi}))^T \mathbf{A} - \mathbf{b}^T\right] \mathbf{J} \end{aligned} \tag{39}$$

Finally, setting the Lagrangian to zero gives the system to solve :

$$\begin{cases} 2\mathbf{J}^T \left[\mathbf{A}\ \text{vec}\,(\mathbf{R}(\boldsymbol{\xi})) - \mathbf{b}\right] + \lambda\boldsymbol{\xi} &= \mathbf{0} \\ \frac{1}{2}\left(x^2+y^2+z^2+w^2-1\right) &= 0 \end{cases} \tag{40}$$

It is a system of polynomial equations of the third degree with five equations and five unknowns including the Lagrange multiplier. This kind of system has no closed form as far as we have investigate. Therefore numerical thechniques must be used to solve it.
It is interesting to note that multiplying the first equation of the system (40) by $\boldsymbol{\xi}^T$ gives an expression for λ :

$$\begin{aligned} \lambda\boldsymbol{\xi}^T\boldsymbol{\xi} &= \boldsymbol{\xi}^T 2\mathbf{J}^T[\mathbf{b} - \mathbf{A}\ \text{vec}\,(\mathbf{R}(\boldsymbol{\xi}))] \\ \lambda &= [2\mathbf{J}\boldsymbol{\xi}]^T[\mathbf{b} - \mathbf{A}\ \text{vec}\,(\mathbf{R}(\boldsymbol{\xi}))] \end{aligned} \tag{41}$$

Substituing eq. (41) into the first equation of system (40) gives a polynomial system of the fifth degree with four equations and four unknowns :

$$\mathbf{J}^T[\mathbf{A}\ \text{vec}\,(\mathbf{R}(\boldsymbol{\xi})) - \mathbf{b}] - (\boldsymbol{\xi}^T\mathbf{J}^T[\mathbf{A}\ \text{vec}\,(\mathbf{R}(\boldsymbol{\xi})) - \mathbf{b}])\ \boldsymbol{\xi} \quad = \quad \mathbf{0}\ . \tag{42}$$

Let

$$\boldsymbol{\phi} \triangleq \mathbf{J}^T[\mathbf{A}\ \text{vec}\,(\mathbf{R}(\boldsymbol{\xi})) - \mathbf{b}] \tag{43}$$

then eq. (42) becomes :

$$\boldsymbol{\phi} - (\boldsymbol{\xi}^T\boldsymbol{\phi})\ \boldsymbol{\xi} \quad = \quad \mathbf{0} \tag{44}$$

where $\boldsymbol{\xi}^T\boldsymbol{\phi}$ has a geometric meaning as the scalar product between $\boldsymbol{\xi}$ and $\boldsymbol{\phi}$.

As $\boldsymbol{\xi}$ has unit norm, a geometric view of the previous equation is that $\boldsymbol{\phi}$ and $\boldsymbol{\xi}$ are parallel and have same orientation.

2.1.1 Numerical approach

We can rewrite this system as $\mathbf{f} = \mathbf{0}$ and the unknowns as the vector $\mathbf{x} = \begin{bmatrix} x & y & z & w & \lambda \end{bmatrix}^T$. Thus the Jacobian of the system, obtained using the chain rule for Hessian matrices, is defined as $\left[\frac{\partial\ \mathbf{f}(i)}{\partial\ \mathbf{x}(j)}\right]_{i,j}$ is :

$$\begin{bmatrix} 4\mathbf{J}^T\mathbf{A}\mathbf{J} + \boldsymbol{\mathcal{G}} + \lambda\mathbf{I}_{4\times 4} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^T & 0 \end{bmatrix} \tag{45}$$

where $\boldsymbol{\mathcal{G}}$ is a 4×4 matrix whose entries are defined by :

$$\boldsymbol{\mathcal{G}}(i,j) \triangleq \left[\ \mathrm{vec}\left(\mathbf{R}(\boldsymbol{\xi})\right)^T \mathbf{A} - \mathbf{b}^T\right] \mathbf{h}_{i,j}$$

and

$$\mathbf{h}_{i,j} \triangleq \frac{\partial\ \ \mathrm{vec}\left(\mathbf{R}\right)\left(\boldsymbol{\xi}\right)}{\partial\ \boldsymbol{\xi}(i)\partial\ \boldsymbol{\xi}(j)}\ .$$

We have $\mathbf{h}_{i,j} = \mathbf{h}_{j,i}$ such that :

$\mathbf{h}_{1,1}$	$\mathbf{h}_{1,2}$	$\mathbf{h}_{1,3}$	$\mathbf{h}_{1,4}$	$\mathbf{h}_{2,2}$	$\mathbf{h}_{2,3}$	$\mathbf{h}_{2,4}$	$\mathbf{h}_{3,3}$	$\mathbf{h}_{3,4}$	$\mathbf{h}_{4,4}$
$\begin{bmatrix}0\\0\\0\\0\\-4\\0\\0\\0\\-4\end{bmatrix}$	$\begin{bmatrix}0\\2\\0\\2\\0\\0\\0\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\2\\0\\0\\0\\2\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\0\\0\\0\\2\\0\\-2\\0\end{bmatrix}$	$\begin{bmatrix}-4\\0\\0\\0\\0\\0\\0\\0\\-4\end{bmatrix}$	$\begin{bmatrix}0\\0\\0\\0\\0\\2\\0\\2\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\-2\\0\\0\\0\\2\\0\\0\end{bmatrix}$	$\begin{bmatrix}-4\\0\\0\\0\\-4\\0\\0\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\2\\0\\-2\\0\\0\\0\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\end{bmatrix}$

gives all the $\mathbf{h}_{i,j}$.

Given a good guess, the system can be solved using an iterative numerical method such as the Newton algorithm.

2.2 Choice of a guess

2.2.1 Ideal case

In the above form, it is clear that the problem has at least one solution. Since the unit quaternions form a sphere which is compact and since Σ^2 is polynomial in $\boldsymbol{\xi}$ and hence continuous, it exists a minimum of the function on the quaternion sphere. It is worthwhile to notice that as Σ^2 is positive, when there is no noise, *i.e.*

$$\mathbf{v}_i = \mathbf{0},\ \forall i$$

in eq. (2), the correct rotation matrix achieves a global minimum among all the matrices too. Since this minimum is unique as the solution of the linear equation $\nabla\Sigma^2(\operatorname{vec}(\mathbf{R})) = \mathbf{0}$, (34), the correct rotation matrix is also given by

$$\operatorname{vec}(\mathbf{R}) = \mathbf{A}^{-1}\mathbf{b} \tag{46}$$

$$\mathbf{R} = \operatorname{mat}_{3\times 3}\left(\mathbf{A}^{-1}\mathbf{b}\right) . \tag{47}$$

2.2.2 Perturbed case

In presence of noise, these relations do not hold anymore. Nevertheless we can use them to define a guess. It is not possible to use $\operatorname{mat}_{3\times 3}\left(\mathbf{A}^{-1}\mathbf{b}\right)$ as a guess in general since it is not even a rotation matrix and hence has no quaternion equivalent. However for a given matrix, we know the closest orthogonal matrix for the Frobenius norm (Schönemann, 1966). Much better in Kanatani (1994) is given an explicit algorithm to find the closest rotation matrix that can be adapted to our case : I therefore propose to use as the guess the quaternion

Algorithm 1 – Closest Rotation Matrix

$\mathbf{A}, \mathbf{b}$ // Inputs

$\mathbf{U}, \mathbf{D}, \mathbf{V} \leftarrow$ the SVD of $\operatorname{mat}_{3\times 3}\left(\mathbf{A}^{-1}\mathbf{b}\right)$ // $\operatorname{mat}_{3\times 3}\left(\mathbf{A}^{-1}\mathbf{b}\right) = \mathbf{U}\mathbf{D}\mathbf{V}^T$ with $\mathbf{D}$ diagonal and $\mathbf{U}$ and $\mathbf{V}$ orthogonal

$\mathbf{R} \leftarrow \mathbf{U}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(\mathbf{U}\mathbf{V}^T) \end{bmatrix}\mathbf{V}^T$

return $\mathbf{R}$

equivalent to the closest rotation matrix, given by the above algorithm.
This choice gives very good results in practice when the noise is reasonable and the matchings are correct.

2.3 Remarks

It is worth noticing that in Viklands (2006, paper1, page 41) a smarter algorithm is proposed that have much stronger convergence properties than the standard Newton minimization. It uses a, so called, *Newton on manifold* procedure that aims to find optimal trajectories on the sphere of rotation matrices. However, that algorithm is also local (it converges also to a local optimum). Moreover, it is slower than the brute force Newton algorithm proposed above for simple cases (when the matchings are almost correct as in a pursuit scenario when initialization is always close to the optimum).

3 Closed form solution in the 2-D case

Theorem (1) is also valid when the space is 2-D. Also the search of the rotation does not require here to involve quaternions. In 2-D, any rotation matrix $\mathbf{R}$ can be written :

$$\mathbf{R} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \quad (48)$$

where α is the real part and β is the imaginary part of the complex number :

$$\boldsymbol{\xi} = \alpha + \mathbf{i}\beta \quad (49)$$

such that $\alpha^2 + \beta^2 = 1$. This constraint is exactly the determinant constraint stated above.

This is now possible to cast the problem into a constrained minimization framework which Lagrangian is :

$$\mathscr{L}_E(\boldsymbol{\xi}, \mathbf{t}) = \Sigma_E^2(\mathbf{R}(\boldsymbol{\xi}), \mathbf{t}) + \lambda(\boldsymbol{\xi}^T\boldsymbol{\xi} - 1) \quad (50)$$

(it has the exact same form than in the 3-D case).

3.1 Minimization

The Jacobian matrix of the change of variable from $\boldsymbol{\xi}$ to $\operatorname{vec}(\mathbf{R})$ is :

$$\mathbf{D}_{\operatorname{vec}(\mathbf{R})(\boldsymbol{\xi})} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (51)$$

Therefore using the chain rule and the Jacobian of the change of variable from $\boldsymbol{\xi}$ to $\operatorname{vec}(\mathbf{R})$ (eq. 51) in 2-D, the system to solve, here, is :

$$\begin{cases} \mathbf{D}^T_{\operatorname{vec}(\mathbf{R})(\alpha,\beta)} \left[\mathbf{A} \operatorname{vec}(\mathbf{R}(\alpha,\beta)) - \mathbf{b}\right] + \lambda \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \mathbf{0} \\ \alpha^2 + \beta^2 - 1 = 0 \end{cases} \quad (52)$$

By expanding all the matrix calculus, it gives :

$$\begin{cases} \alpha(x_1 + \lambda) + \beta y_1 - z_1 = 0 \\ \alpha y_1 + \beta(y_2 + \lambda) - z_2 = 0 \\ \alpha^2 + \beta^2 = 1 \end{cases} \quad (53)$$

where

$$\begin{aligned} x_1 &= \mathbf{A}(1,1) + \mathbf{A}(4,1) + \mathbf{A}(1,4) + \mathbf{A}(4,4)\ , \\ y_1 &= \mathbf{A}(2,1) - \mathbf{A}(3,1) + \mathbf{A}(2,4) - \mathbf{A}(3,4)\ , \\ z_1 &= \mathbf{b}(1) + \mathbf{b}(4)\ , \\ y_2 &= \mathbf{A}(2,2) - \mathbf{A}(3,2) - \mathbf{A}(2,3) + \mathbf{A}(3,3)\ , \\ z_2 &= \mathbf{b}(2) - \mathbf{b}(3)\ . \end{aligned}$$

Now it is possible to obtain $\begin{bmatrix} \alpha & \beta \end{bmatrix}^T$ as a function of λ :

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{(x_1 + \lambda)(y_2 + \lambda) - y_1^2} \begin{bmatrix} -z_2 y_1 + z_1(y_2 + \lambda) \\ z_2(x_1 + \lambda) - z_1 y_1 \end{bmatrix} . \quad (54)$$

With that system of equations, it is clear that $z_1 = z_2 = 0$ implies $\alpha = \beta = 0$ and indeed the problem has no solution. By remplacing α and β into the last equation of the system (53) we obtain :

$$\begin{aligned} P(\lambda) = \lambda^4 + 2\delta\lambda^3 + \left[\ \delta^2 - 2\Delta_3 - z_1^2 - z_2^2\right]\lambda^2 \\ +2(z_1\Delta_1 + z_2\Delta_2 - \delta\Delta_3)\lambda + \Delta_3^2 - \Delta_1^2 - \Delta_2^2 \quad = \quad 0\ , \end{aligned} \tag{55}$$

where

$$\begin{aligned} \delta &= x_1 + y_2\ , \\ \Delta_1 &= z_2 y_1 - z_1 y_2\ , \\ \Delta_2 &= z_1 y_1 - z_2 x_1\ , \\ \Delta_3 &= y_1^2 - y_2 x_1\ . \end{aligned}$$

$P(\lambda)$ is a 4 degree polynomial with respect to λ. The problem is solved by finding the roots of P which agrees with the result of Censi (2008). λ is necessarily real thus the searched Lagrange multiplier is found by taking the real root that minimizes Σ^2.

Theorem 3 (Choice of the optimal Lagrange multiplier in the 2-D case) *It is not necessary to compute the energy associated with a Lagrange multiplier to obtain the global minimizer because the optimal Lagrange multiplier is always given by the greatest Lagrange multiplier. Let* $(\mathbf{R}_1, \lambda_1)$ *and* $(\mathbf{R}_2, \lambda_2)$ *be two solutions of the system (52), the link with the associated energy values* $\Sigma^2(\mathbf{R}_1)$ *and* $\Sigma^2(\mathbf{R}_2)$ *is given by the following inequality :*
if $\lambda_1 \geq \lambda_2$ *:*

$$\frac{1}{2}[\lambda_1 - \lambda_2] \leq \Sigma^2(\mathbf{R}_2) - \Sigma^2(\mathbf{R}_1) \tag{56}$$

Proof:

The steps of the proof given here are inspired by the proof of theorem (1) in Gander (1980), unfortunately this theorem does not apply here, even if the problem studied in that work is very close to our case. As $\mathbf{D}_{\ \mathrm{vec}(\mathbf{R})(\alpha,\beta)}$ is constant let us note it $\mathbf{D}$ in this proof. The main observation is that :

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix}^T \mathbf{D}^T \quad = \quad \begin{bmatrix}\alpha\\ \beta\\ -\beta\\ \alpha\end{bmatrix} = \ \mathrm{vec}\left(\mathbf{R}(\alpha,\beta)\right)^T \ . \tag{57}$$

Let us note now, $\mathbf{x}_i$ for $\mathrm{vec}\left(\mathbf{R}(\alpha_i,\beta_i)\right)$, $i = 1, 2$. If $\mathbf{x}_1, \lambda_1$ and $\mathbf{x}_2, \lambda_2$ are solution of the system (52), then :

$$\mathbf{D}^T[\mathbf{A}\mathbf{x}_1 - \mathbf{b}] + \lambda_1 \begin{bmatrix}\alpha_1\\ \beta_1\end{bmatrix} \quad = \quad \mathbf{0}\ , \tag{58}$$

$$\mathbf{D}^T[\mathbf{A}\mathbf{x}_2 - \mathbf{b}] + \lambda_2 \begin{bmatrix}\alpha_2\\ \beta_2\end{bmatrix} \quad = \quad \mathbf{0}\ . \tag{59}$$

Multiplying eq. (58) by $\begin{bmatrix}\alpha_1\\\beta_1\end{bmatrix}^T$ and eq. (59) by $\begin{bmatrix}\alpha_2\\\beta_2\end{bmatrix}^T$ to the left, one obtains using the observation (57) :

$$\begin{aligned}\mathbf{x}_1^T[\mathbf{A}\mathbf{x}_1 - \mathbf{b}] + \lambda_1(\alpha_1^2 + \beta_1^2) &= 0\ ,\\ \mathbf{x}_2^T[\mathbf{A}\mathbf{x}_2 - \mathbf{b}] + \lambda_2(\alpha_2^2 + \beta_2^2) &= 0\ .\end{aligned}$$

Using the norm constraint :

$$\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 - \mathbf{b}^T\mathbf{x}_1 + \lambda_1 = 0\ , \tag{60}$$
$$\mathbf{x}_2^T\mathbf{A}\mathbf{x}_2 - \mathbf{b}^T\mathbf{x}_2 + \lambda_2 = 0\ . \tag{61}$$

Now substracting eq. (61) from eq. (60) gives :

$$\lambda_1 - \lambda_2 = (\mathbf{x_2}^T\mathbf{A}\mathbf{x}_2 - \mathbf{b}^T\mathbf{x_2}) - (\mathbf{x_1}^T\mathbf{A}\mathbf{x}_1 - \mathbf{b}^T\mathbf{x_1}) \tag{62}$$

Multiplying eq. (58) by $\begin{bmatrix}\alpha_2\\\beta_2\end{bmatrix}^T$ and eq. (59) by $\begin{bmatrix}\alpha_1\\\beta_1\end{bmatrix}^T$ to the left, one obtains using the observation (57) :

$$\mathbf{x}_2^T\mathbf{A}\mathbf{x}_1 - \mathbf{b}^T\mathbf{x}_2 + \lambda_1\frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2 = 0\ , \tag{63}$$
$$\mathbf{x}_1^T\mathbf{A}\mathbf{x}_2 - \mathbf{b}^T\mathbf{x}_1 + \lambda_2\frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2 = 0\ , \tag{64}$$

noticing that $\begin{bmatrix}\alpha_1\\\beta_1\end{bmatrix}^T\begin{bmatrix}\alpha_2\\\beta_2\end{bmatrix} = \frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2$ Now substracting eq. (64) from eq. (63) gives :

$$-\frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2(\lambda_1 - \lambda_2) = \mathbf{b}^T\mathbf{x}_1 - \mathbf{b}^T\mathbf{x}_2 \tag{65}$$

using the equality $\mathbf{x}_2^T\mathbf{A}\mathbf{x}_1 = \mathbf{x}_1^T\mathbf{A}\mathbf{x}_2$ since $\mathbf{A}$ is symmetric. Now if we sum eq. (62) and eq. (65), we obtain :

$$(\lambda_1 - \lambda_2) - \frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2(\lambda_1 - \lambda_2) = (\mathbf{x}_2^T\mathbf{A}\mathbf{x}_2 - 2\mathbf{b}^T\mathbf{x}_2) - (\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 - 2\mathbf{b}^T\mathbf{x}_1) \tag{66}$$
$$(\lambda_1 - \lambda_2)[1 - \frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2] = (\mathbf{x}_2^T\mathbf{A}\mathbf{x}_2 - 2\mathbf{b}^T\mathbf{x}_2 + \kappa) - (\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 - 2\mathbf{b}^T\mathbf{x}_1 + \kappa)$$
$$(\lambda_1 - \lambda_2)[1 - \frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2] = \Sigma^2(\mathbf{x}_2) - \Sigma^2(\mathbf{x}_1) \tag{67}$$

Using Cauchy-Schwartz inequality we have :

$$\begin{aligned}\mathbf{x}_1^T\mathbf{x}_2 &\le \|\mathbf{x}_1\|\|\mathbf{x}_2\|\\ \mathbf{x}_1^T\mathbf{x}_2 &\le 1\\ 1 - \frac{1}{2}\mathbf{x}_1^T\mathbf{x}_2 &\ge \frac{1}{2}\end{aligned} \tag{68}$$

Therefore eq. (67) and eq. (68) implies that $\lambda_1 - \lambda_2$ and $\Sigma^2(\mathbf{x}_2) - \Sigma^2(\mathbf{x}_1)$ have the same sign. Therefore if $\lambda_1 \ge \lambda_2$ then $\Sigma^2(\mathbf{x}_1) \le \Sigma^2(\mathbf{x}_2)$ and the solution

$(\mathbf{x}_1, \lambda_1)$ is a better minimizer.
Moreover we have if $\lambda_1 \geq \lambda_2$, then

$$\Sigma^2(\mathbf{x}_2) - \Sigma^2(\mathbf{x}_1) \geq \frac{1}{2}[\lambda_1 - \lambda_2] \ . \tag{56}$$

4 Optimality condition for the 3D case

The complete study of the 2-D case led me to investigate more closely the 3-D case and fortunately, I found that a similar result holds.

Theorem 4 (Choice of the optimal Lagrange multiplier in the 3-D case) *The 3-D case is similar to the 2-D case : it is not necessary to compute the energy associated with a Lagrange multiplier to obtain the global minimizer because the optimal Lagrange multiplier is always given by the greatest Lagrange multiplier.*

Proof:

The main observation is, here, that :

$$\text{vec}\,(\mathbf{R}(\xi)) - \text{vec}\,(\mathbf{I}_{3\times 3}) = \mathbf{J}\boldsymbol{\xi} \tag{69}$$

Let us note now, $\mathbf{x}_i$ for $\text{vec}\,(\mathbf{R}(\xi_i))$, $i = 1, 2$. If $\mathbf{x}_1, \lambda_1$ and $\mathbf{x}_2, \lambda_2$ are solution of the system (40) :

$$2\mathbf{J}^T[\mathbf{A}\mathbf{x}_1 - \mathbf{b}] + \lambda_1\boldsymbol{\xi}_1 = \mathbf{0} \ , \tag{70}$$
$$2\mathbf{J}^T[\mathbf{A}\mathbf{x}_2 - \mathbf{b}] + \lambda_2\boldsymbol{\xi}_2 = \mathbf{0} \ . \tag{71}$$

thus, following closely the steps of the proof of the theorem (3), multiplying each equation (70) and (71) by $\boldsymbol{\xi}_1^T$ and $\boldsymbol{\xi}_2^T$ in that order then in cross order ; then substracting gives :

$$\begin{aligned}\lambda_1 - \lambda_2 &= 2(\mathbf{x_2}^T\mathbf{A}\mathbf{x}_2 - \mathbf{b}^T\mathbf{x_2} - \text{vec}\,(\mathbf{I}_{3\times 3})^T\,\mathbf{A}\mathbf{x}_2) \\ &\quad -2(\mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 - \mathbf{b}^T\mathbf{x}_2 - \text{vec}\,(\mathbf{I}_{3\times 3})^T\,\mathbf{A}\mathbf{x}_1) \qquad (72)\\ -\mathbf{x}_1^T\mathbf{x}_2(\lambda_1 - \lambda_2) &= 2(\mathbf{b}^T\mathbf{x}_1 - \mathbf{b}^T\mathbf{x}_2) \\ &\quad -2(\text{vec}\,(\mathbf{I}_{3\times 3})^T\,\mathbf{A}\mathbf{x}_1 - \text{vec}\,(\mathbf{I}_{3\times 3})^T\,\mathbf{A}\mathbf{x}_2) \qquad (73)\end{aligned}$$

The next step is to sum eq. (72) and eq. (73) :

$$\begin{aligned}(1 - \mathbf{x}_1^T \mathbf{x}_2)(\lambda_1 - \lambda_2) &= 2[(\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2 - 2\mathbf{b}^T \mathbf{x}_2) - (\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 - 2\mathbf{b}^T \mathbf{x_1})] \\ (1 - \mathbf{x}_1^T \mathbf{x}_2)(\lambda_1 - \lambda_2) &= 2[\Sigma^2(\mathbf{x}_2) - \Sigma^2(\mathbf{x}_1)] \qquad (74)\end{aligned}$$

Again we have, thanks to the Cauchy-Schwartz inequality, a usefull inequality : $1 - \mathbf{x}_1^T \mathbf{x}_2 \geq 0$. Now we can conclude that $\lambda_1 - \lambda_2$ and $\Sigma^2(\mathbf{x}_2) - \Sigma^2(\mathbf{x}_1)$ have the same sign.

$\blacksquare$

5 Appendix A : linear algebra results

5.1 Kronecker product

Definition 1 (Kronecker product) *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix and* $\mathbf{B}$ *a* $p \times q$ *matrix. The* $mp \times nq$ *matrix defined by*

$$\begin{bmatrix} \mathbf{A}(1,1)\ \mathbf{B} & \dots & \mathbf{A}(1,n)\ \mathbf{B} \\ \vdots & & \vdots \\ \mathbf{A}(m,1)\ \mathbf{B} & \dots & \mathbf{A}(m,n)\ \mathbf{B} \end{bmatrix}$$

is called the Kronecker product *of* $\mathbf{A}$ *and* $\mathbf{B}$ *and written* $\mathbf{A} \otimes \mathbf{B}$.

The Kronecker product is defined for any pair of matrices $\mathbf{A}$ and $\mathbf{B}$ whenever their sizes are.
For instance if $\mathbf{A} = \begin{bmatrix} a & b & c \end{bmatrix}$ and $\mathbf{B} = \mathbf{I}_{3\times3}$ then

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a\mathbf{I}_{3\times3} & b\mathbf{I}_{3\times3} & c\mathbf{I}_{3\times3} \end{bmatrix} = \begin{bmatrix} a & 0 & 0 & b & 0 & 0 & c & 0 & 0 \\ 0 & a & 0 & 0 & b & 0 & 0 & c & 0 \\ 0 & 0 & a & 0 & 0 & b & 0 & 0 & c \end{bmatrix}$$

Theorem 5 (Transpose and Kronecker product) *Let* $\mathbf{A}$ *be a* $m \times n$ *matrix and* $\mathbf{B}$ *be a* $n \times q$ *matrix then :*

$$[\mathbf{A} \otimes \mathbf{B}]^T = \mathbf{A}^T \otimes \mathbf{B}^T$$

5.2 Vec operator

Definition 2 (Vec operator) *Let* $\mathbf{A}$ *be a* $m \times n$ *matrix and* $\mathbf{a}_j$ *its j-th column ; then* $vec\,(\mathbf{A})$ *is the* $mn \times 1$ *vector :*

$$vec\,(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$$

The operator transforms a matrix into a vector by stacking the columns of the matrix one underneath the other.

There is an obvious, but usefull relation : when $\mathbf{p}$ is a vector :

$$\mathrm{vec}\,(\mathbf{p}) = \mathbf{p}. \tag{75}$$

5.3 Mat operator

Definition 3 (Mat operator) *The mat operator is used to inverse the vec operator and restore a matrix from a vector. However it is necessary to define the dimensions of the output matrix, since different matrices have the same image by the vec operator. Therefore if* $\mathbf{v}$ *is a* n *dimensional vector such that* $n = pq$

$$mat_{p\times q}(\mathbf{v}) = \mathbf{M}$$

where $\mathbf{M}$ *is a* $p\times q$ *matrix such that* $vec(\mathbf{M}) = \mathbf{v}$.

5.4 Relations between the vec operator and the Kronecker product

Theorem 6 (Vectorialization of the product of 3 matrices) *Let* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *three matrices such that the product* $\mathbf{ABC}$ *is defined. Then*

$$vec(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right) vec(\mathbf{B})$$

Theorem 7 (Vectorialization of a matrix product) *Let* $\mathbf{A}$ *be a* $m\times n$ *matrix and* $\mathbf{B}$ *be a* $n\times q$ *matrix then :*

$$vec(\mathbf{AB}) = \left(\mathbf{B}^T \otimes \mathbf{I}_{m\times m}\right) vec(\mathbf{A}) = \left(\mathbf{B}^T \otimes \mathbf{A}\right) vec\left(\mathbf{I}_{n\times n}\right) = \left(\mathbf{I}_{q\times q} \otimes \mathbf{A}\right) vec(\mathbf{B})$$

Table des matières

1 Estimating the optimal rigid body transformation using Mahalanobis distances 3

1.1 Problem setting 3
1.2 $\mathbf{t}^*$ as a function of $\mathbf{R}^*$ 4
1.3 Expansion of Σ^2 4
1.3.1 Expansion of $\mathbf{x}_i^T \mathbf{\Pi}_i \mathbf{x}_i$ 5
1.3.2 Sum over i 6

2 3-D Minimization 9

2.1 Equivalence with a polynomial system 9
2.1.1 Numerical approach 11
2.2 Choice of a guess 11
2.2.1 Ideal case 11
2.2.2 Perturbed case 12
2.3 Remarks 12

3 Closed form solution in the 2-D case 12

3.1 Minimization 13

4 Optimality condition for the 3D case 16

5 Appendix A : linear algebra results 18

5.1 Kronecker product 18
5.2 Vec operator 18
5.3 Mat operator 19
5.4 Relations between the vec operator and the Kronecker product . 19

Références

A. Censi. An ICP variant using a point-to-line metric. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Pasadena, CA, May 2008. URL `http://purl.org/censi/2007/plicp`.

W. Gander. Least squares with a quadratic constraint. *Numerische Mathematik*, 36(3) :291–307, September 1980.

K. Kanatani. Analysis of 3-d rotation fitting. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 16(5) :543–549, 1994. ISSN 0162-8828. doi : http://dx.doi.org/10.1109/34.291441.

P. H. Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31 :1–10, 1966.

T. Viklands. *Algorithms for Weighted Orthogonal Procrustes Problem and other Least Squares Problems.* PhD thesis, Department of Computing Science, Umeå university, Umeå, Sweden, 2006.

Centre de recherche INRIA Grenoble – Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399