Robust Hand-Eye Calibration for Computer Aided Medical Endoscopy

Abed Malti and João P. Barreto
Institute of Systems and Robotics,
Faculty of Science and Technology,
University of Coimbra.
3030 Coimbra, Portugal
{amalti,jpbar}@isr.uc.pt

***Abstract*— Endoscopic camera for surgical navigation and 3D visualization requires precise and stable estimates of the calibration parameters. The estimation of the hand-eye transform between the camera frame and the opto-tracked body of the endoscope is an important issue of the calibration. This paper presents a new stable method for the hand-eye calibration problem. The most popular method estimates the transform directly in the special euclidean group $SE(3)$ by computing separately the rotation and the translation. The second famous approach formulates the problem in the dual quaternion space and estimates jointly the rotation and the translation. In a first glance, the simultaneous estimation seems to be always advantageous. However, and according to the experiments, this is not the case for the rotation estimation that is affected by the noise in translation. Our approach takes advantage of the both methods and uses the dual quaternion to estimate separately the rotation and the translation. We show experimentally that our algorithm is more stable with minimal number and small amplitude of motions.** [1]

I. INTRODUCTION

Developing a vision system for minimal invasive computer aided surgery requires precise calibration of the endoscopic camera. Besides knowing the intrinsics [2], we further constrain the problem by tracking the poses of the camera body using an opto-sensor. However, in order to register the pre-operative models of organs we need to recover the 3D motion of the camera reference frame. Since the camera body and the camera reference frame are not coincident we need to estimate the rigid displacement between them.

Determining such displacement is known as the hand-eye calibration, see Fig. (1). This problem initially arose in robotics when a camera was mounted on a robot to measure 2D and 3D geometric relationships among different viewed objects [13]. It has been applied in several contexts : for instance, in sensor based motion planning [14] in order to automatically determine the optimum positions of the sensor so that all the desired features can be viewed while taking care of problems of occlusion, depth of focus, field of view, etc. We can cite also in grasping an object [1], where even if the vision system is able to determine the relative pose of the object to the camera, it is necessary that the robot knows how to place the manipulator to grasp it.

[1]The authors acknowledge the Portuguese Science Foundation, that generously funded this work through grant PTDC/EEA-ACR/68887/2006.

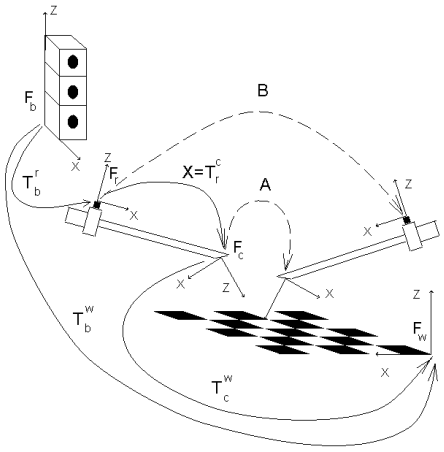


Fig. 1. The Hand-Eye problem. In our case, the body of the endoscope represents the effector in the classic formulation. Its frame F_r is attached to a marker. The pose of the marker is tracked by an opto-sensor. Without loss of generality, we can assume that the world reference frame is attached to the calibration chess-grid. The camera calibration enables to estimate the pose T_c^w between the camera and the world reference frame. For two different poses at the instants τ and τ', the hand-eye transform T_r^c is the solution of the equation $A\,X = X\,B$. Where $A = T_w^c(\tau)\,T_c^w(\tau')$ represents the motion of the camera reference frame and $B = T_b^r(\tau)\,T_r^b(\tau')$ encodes the endoscope motion tracked by the opto-sensor.

Mathematically speaking, the hand-eye calibration consists on solving for $X \in SE(3)$ the following equation :

$$A\,X = X\,B \tag{1}$$

Where $A \in SE(3)$ represents the motion of the camera reference frame and $B \in SE(3)$ encodes the motion of the rigid body to which the camera is attached. Tsai and Lenz [11] prove that at least 2 motions with different rotation axis are needed to solve this equation. It can be observed that the rotation involved by the camera motion A is the same as the one involved by the effector motion B. This intrinsic property of the problem is highly relevant to balance the estimation of the rotation part of X in presence of noise.

Several methods exist in the literature to solve this equation. The different approaches can be divided in three groups : **(i)** Separated estimation in $SE(3)$: It decomposes the Eq.1 into a rotational and a translational part and solves for first the rotation and then the translation [11], [15], [4], [5]. **(ii)** Simultaneous estimation : The idea is to

simultaneously solve for the translation and rotation using a screw representation of the problem. The most famous method is proposed by [6] where dual quaternions are used to encode the screw representation of the hand-eye problem. **(iii)** Estimation including camera intrinsics : The hand-eye transform and the camera intrinsics are simultaneously estimated in a global iterative optimization step [7]. In this paper we assume that the camera intrinsics have been previously estimated [2]. Therefore, we will not consider this method in the comparative analysis.

The hand-eye calibration applied to the context of endoscopy has specificities: (i) the translation components of the motions A and B are usually small (below 10 cm) due to the close range image characteristics of the device, (ii) the calibration has to be performed by a non-expert in the operating room [12], which requires the method to be robust with a minimum number of motions.

We ran several experiments and observed that under such considerations, neither the separated estimation in $SE(3)$, nor the joint estimation using dual quaternions, lead to a satisfactory result. The first method provides good estimates for the rotation. However the results for translation are rather inaccurate because of the small motions. In the second method, the rotation enables to constraint the translation estimate leading to better results in translation. Unfortunately, and since rotation and translation are determined at the same time, the latter seems to have a negative impact in the former, and the results for rotation are more inaccurate than the ones obtained with the first method.

The main contribution of this work is a robust hand-eye calibration algorithm with small amplitude and minimal number of motions. The proposed method overcomes the detected difficulties by taking the advantages of both classic methods. It uses the dual quaternion formulation of the problem because it offers a stable representation of the translation as a shifting along the rotation axis. It estimates separately the rotation and the translation for mainly two reasons : **(1)** The rotation part is fully self-characterized through the trace property. **(2)** The noise in translation badly affects the rotation in a joint estimation. The method, although designed for hand-eye calibration in medical endoscopy, can be successfully applied in other contexts of hand-eye calibration.

A. Structure and Notation

This paper is organized as follows : Section II introduces a brief review about quaternions and dual quaternions. Section III presents the separated estimation in $SE(3)$ using quaternions. Section IV states the formulation of the problem in the dual quaternion space and gives a sight of the classic approach solution. Section V presents the proposed approach. Finally, sections VI and VII present synthetic and real experiments within discussion and comments.

Quaternions are represented by lowercase bold font (e.g. $\mathbf{q}$). Matrices are denoted by uppercase sans serif font (e.g. A). Vectors providing a direction in 3D are represented using plain lowercase topped by an arrow (e.g. $\vec{l}$). For convenience, and given two 3×1 vectors $\vec{l}$ and $\vec{m}$, the dot product is indicated either using $< .\,,\, . >$ or using regular matrix/vector multiplication (e.g. $< \vec{l}, \vec{m} >= \vec{l}^{\mathsf{T}}\,\vec{m}$) and the cross product is carried either using the symbol $\times$ or using the skew symmetric matrix (e.g. $\vec{l} \times \vec{m} = [\vec{l}]_{\times}\,\vec{m}$). For sake of simplicity, $||.||_2$ denotes the vector norm 2 either in $\mathbb{R}^3$ or in $\mathbb{R}^4$. In both cases, the concerned space will be mentioned.

II. Background

A. Representation of rotations using quaternions

A quaternion $\mathbf{q} \in \mathbb{Q}$ is a quadruplet of real numbers [9] that can be split in a scalar q_0 and a 3D vector component $\vec{q}$:

$$\mathbf{q} = \begin{pmatrix} q_0 \\ \vec{q} \end{pmatrix} \tag{2}$$

The conjugate quaternion $\mathbf{q}^*$ of $\mathbf{q}$ is defined as :

$$\mathbf{q}^* = \begin{pmatrix} q_0 \\ -\vec{q} \end{pmatrix} \tag{3}$$

The product of two quaternions a and b is defined as follows :

$$\mathbf{a}\,.\,\mathbf{b} = \begin{pmatrix} a_0\, b_0 - \vec{a}^{\mathsf{T}}.\vec{b} \\ \\ a_0\, \vec{b} + b_0\, \vec{a} + \vec{a} \times \vec{b} \end{pmatrix} \tag{4}$$

A quaternion q represents uniquely a rotation matrix $\mathsf{R} \in SO(3)$ if and only if it is a unit quaternion :

$$\mathbf{q}^* \cdot \mathbf{q} = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} \tag{5}$$

This condition encodes the orthogonal constraint fulfilled by the $SO(3)$ matrices. Since a rotation matrix has 3 degrees of freedom (DOF), this constraint defines a 3 dimensional sub-manifold in the 4D quaternion space. The one-to-one mapping from $SO(3)$ to this sub manifold is given by :

$$\mathbf{q} = \begin{pmatrix} \sin(\frac{\theta}{2}) \\ \cos(\frac{\theta}{2})\,\vec{l} \end{pmatrix} \tag{6}$$

Where $\vec{l} \in \mathbb{R}^3$ (with $||\vec{l}|| = 1$) is the rotation axis and $\theta \in [-\pi, \pi]$ is the rotation angle. The composition product of rotations is represented as a quaternion product and the inverse rotation is represented as the conjugate quaternion.

B. Representation of rigid motions using dual quaternions

Consider a rigid motion in $SE(3)$ represented by a 4×4 matrix T with 6 DOF (3 DOF for the rotation R and 3 DOF for the translation $\vec{t}$) :

$$\mathsf{T} = \begin{pmatrix} \mathsf{R} & \vec{t} \\ 0 & 1 \end{pmatrix} \tag{7}$$

Any rigid motion T can be carried by assuming a rotation of an angle θ around a 3D line $\mathbf{s}$, and a translation d along this same line (c.f. *Chasles theorem*[10]). This leads to the screw representation for rigid transformations consisting in a line in 3D space (represented by a 6×1 vector $\mathbf{s}$), a rotation

angle θ, and a pitch value d. For a particular $\mathsf{T} \in SE(3)$ the screw axis is :

$$\mathbf{s} = \begin{pmatrix} \vec{l} \\ \vec{m} \end{pmatrix} = \begin{pmatrix} \vec{l} \\ \frac{1}{2}\,[\vec{t}]_\times \vec{l} + \frac{1}{2}\cot(\frac{\theta}{2})[\vec{l}]_\times [\vec{t}]_\times \vec{l} \end{pmatrix} \tag{8}$$

with $\vec{t}$ being the translation component, $\vec{l}$ and θ being respectively the axis and the angle of the rotation R. Remark that $\vec{l}$ and the momentum vector $\vec{m}$ are always orthogonal to each other. The pitch d is given by :

$$d = < \vec{t}, \vec{l} > \tag{9}$$

In the same manner that a rotation R can be represented by a quaternion $\mathbf{q}$, a rigid displacement T can be described using a dual quaternion $\widehat{\mathbf{q}}$. A dual quaternion has the following form:

$$\widehat{\mathbf{q}} = \mathbf{q} + \epsilon\,\mathbf{q}' = \begin{pmatrix} q_0 \\ \vec{q} \end{pmatrix} + \epsilon \begin{pmatrix} q_0' \\ \vec{q}' \end{pmatrix} \tag{10}$$

with $\mathbf{q}$ and $\mathbf{q}'$ being quaternions and ϵ being a scalar constant such that $\epsilon^2 = 0$. $\mathbf{q}$ and $\mathbf{q}'$ are usually referred as the real and the dual components. The conjugate of a dual quaternion $\widehat{\mathbf{q}}$ is defined as :

$$\widehat{\mathbf{q}}^* = \mathbf{q}^* + \epsilon\,\mathbf{q}'^* \tag{11}$$

The product of two dual quaternions $\widehat{\mathbf{a}}$ and $\widehat{\mathbf{b}}$ is carried as follows :

$$\widehat{\mathbf{a}}.\widehat{\mathbf{b}} = (\mathbf{a}.\mathbf{b}) + \epsilon(\mathbf{a}.\mathbf{b}' + \mathbf{a}'.\mathbf{b}) \tag{12}$$

To represent a $SE(3)$ element, $\widehat{\mathbf{q}}$ has to be a unit dual quaternion :

$$\widehat{\mathbf{q}}^* \,.\, \widehat{\mathbf{q}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{13}$$

In other terms, $\mathbf{q}$ has to be a unit quaternion and has to verify an orthogonality condition with $\mathbf{q}'$:

$$q_0\, q_0' + < \vec{q}, \vec{q}' >= 0\,, \tag{14}$$

Consider the motion T and its screw representation discussed above. T can be represented by a dual quaternion $\widehat{\mathbf{q}}$ where the real component $\mathbf{q}$ is the quaternion corresponding to the rotation R, and where the dual part $\mathbf{q}'$ is :

$$\mathbf{q}' = \begin{pmatrix} -\frac{d}{2}\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2})\vec{m} + \frac{d}{2}\cos(\frac{\theta}{2})\vec{l} \end{pmatrix} \tag{15}$$

Let A and B be two rigid transformations. The dual quaternion representation of $\mathsf{T} = \mathsf{A}\,\mathsf{B}$ is $\widehat{\mathbf{q}} = \widehat{\mathbf{a}}.\widehat{\mathbf{b}}$.

Let $\widehat{\mathbf{q}}$ be a dual quaternion representing a rigid motion. Since the real part $\mathbf{q}$ is the quaternion encoding the rotation, recovering matrix R is trivial. To determine the translational component of the motion we can use the following relation involving the conjugate of $\mathbf{q}$:

$$\begin{pmatrix} 0 \\ \vec{t} \end{pmatrix} = 2\,\mathbf{q}'.\mathbf{q}^* \tag{16}$$

III. Classic solving in SE(3)

For one motion of the camera body, the hand-eye formulation of Eq.(1) can be rewritten as :

$$\begin{pmatrix} \mathsf{R}_A & \vec{t}_A \\ 0_{3\times3} & 1 \end{pmatrix} \begin{pmatrix} \mathsf{R}_X & \vec{t}_X \\ 0_{3\times3} & 1 \end{pmatrix} = \begin{pmatrix} \mathsf{R}_X & \vec{t}_X \\ 0_{3\times3} & 1 \end{pmatrix} \begin{pmatrix} \mathsf{R}_B & \vec{t}_B \\ 0_{3\times3} & 1 \end{pmatrix} \tag{17}$$

A. Solving for rotation

Considering only the rotation part, it comes that :

$$\mathsf{R}_A\,\mathsf{R}_X \;=\; \mathsf{R}_X\,\mathsf{R}_B \tag{18}$$

Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{q}$ be the quaternions associated with R_A, R_B and R respectively. Using the quaternion multiplication it follows :

$$\mathbf{a}.\mathbf{q} \;=\; \mathbf{q}.\mathbf{b} \tag{19}$$

Using the result of Eq.(4), and performing some algebraic operations on Eq. (19) we conclude that :

$$\underbrace{\begin{pmatrix} a_0 - b_0 & -(\vec{a} - \vec{b})^{\mathsf{T}} \\ \vec{a} - \vec{b} & [\vec{a} + \vec{b}]_\times + (a_0 - b_0)\mathsf{I}_3 \end{pmatrix}}_{\mathsf{K}(\mathbf{a},\mathbf{b})} \mathbf{q} \;=\; 0 \tag{20}$$

With I_3 being the 3×3 identity matrix. The quaternion equation can be written in matrix form as :

$$\mathsf{K}(\mathbf{a},\mathbf{b})\,\mathbf{q} \;=\; 0 \tag{21}$$

Where $\mathsf{K}(\mathbf{a},\mathbf{b})$ is a 4×4 matrix defined for each pair of quaternions $(\mathbf{a},\mathbf{b})$.

Equation (18) is a particular case of a similarity transformation. It is well known that it exists a solution X *iff* the trace of R_A is equal to the trace of R_B (rotations with the same angle θ but different axis). This means that the scalar part of $\mathbf{a}$ and $\mathbf{b}$ are equal ($a_0 = b_0$), and then the vector parts have the same modulus ($||\vec{a}|| = ||\vec{b}||$).

In general the Eq.(21) only admits the trivial solution because $\mathsf{K}(.,.)$ is full rank. For the case of $\mathbf{a}$ and $\mathbf{b}$ being two rotations with the same angle, matrix $\mathsf{K}(.,.)$ becomes of rank 3 and there is a valid quaternion solution $\mathbf{q}$ different from zero. The existence of a solution can be understood geometrically. Since ($a_0 = b_0$) and ($||\vec{a}|| = ||\vec{b}||$) then $(\vec{a} + \vec{b})$ is always orthogonal to $(\vec{a} - \vec{b})$. Thus, according to the first line of the matrix $\mathsf{K}(.,.)$, it comes out that the vector component of the solution $\mathbf{q}$ must be orthogonal to both $\vec{a}$ and $\vec{b}$. It can be verified that this information is already encoded in the bottom 3-lines bloc of this matrix. Indeed, it states that $(\vec{a} - \vec{b})$ and $((\vec{a} + \vec{b}) \times \vec{q})$ must be collinear which is possible *iff* $\vec{q}$ is orthogonal to both vectors $(\vec{a} - \vec{b})$ and $(\vec{a} + \vec{b})$. These statements allow to discard redundant terms in matrix $K(.,.)$ that can be rewritten as :

$$\tilde{\mathsf{K}}(\mathbf{a},\mathbf{b}) = \begin{pmatrix} \vec{a} - \vec{b} & [\vec{a} + \vec{b}]_\times \end{pmatrix} \tag{22}$$

Where $\tilde{\mathsf{K}}(\mathbf{a},\mathbf{b})$ is a 3×4 matrix defined for each pair of quaternions $(\mathbf{a},\mathbf{b})$.

In practice the rotation R_A and R_B are measured and, due to the noise, they do not have the same trace. Hence, the full formulation of Eq.(21) is important to balance the scalar condition in presence of noise. The solution $\mathbf{q}$ for the rotation is determined in the least square sense by the SVD decomposition of

$$\mathsf{L} = \begin{pmatrix} \mathsf{K}(\mathbf{a}_1,\mathbf{b}_1) \\ \vdots \\ \mathsf{K}(\mathbf{a}_N,\mathbf{b}_N) \end{pmatrix} \tag{23}$$

with $\mathsf{K}(\mathbf{a}_i, \mathbf{b}_i)$, $i = 1, 2 \ldots, N$, being the matrices of the N considered motions $\{(\mathsf{A}_1, \mathsf{B}_1), \ldots, (\mathsf{A}_N, \mathsf{B}_N)\}$. Theoretically, only one motion is enough to solve the rotation part, which is not the case for the translation.

B. Solving for translation

The translation is determined after getting the rotation. Considering the translation part of Eq.(17), it comes that :

$$(\mathsf{R}_A - \mathsf{I}_3)\,\vec{t}_X = \mathsf{R}_X\,\vec{t}_B \;-\; \vec{t}_A \tag{24}$$

The matrix $(\mathsf{R}_A - \mathsf{I}_3)$ is of rank 2 because $((\mathsf{R}_A - \mathsf{I}_3)\,\vec{l}_A = 0)$, with $\vec{l}_A$ being the rotation vector associated to R_A. Henceforth, at least $N = 2$ motions are needed to solve for the 3D translation vector. For $N \geq 2$ motions, the solution for the translation $\vec{t}_X$ is formulated as a non-constrained least square problem :

$$\min \left\| \underbrace{\begin{pmatrix} \mathsf{R}_{A_1} - \mathsf{I}_3 \\ \vdots \\ \mathsf{R}_{A_N} - \mathsf{I}_3 \end{pmatrix}}_{\mathsf{D}} \vec{t}_x \;-\; \begin{pmatrix} \mathsf{R}_X\,\vec{t}_{B_1} \;-\; \vec{t}_{A_1} \\ \vdots \\ \mathsf{R}_X\,\vec{t}_{B_N} \;-\; \vec{t}_{A_N} \end{pmatrix} \right\|_2 \tag{25}$$

With $||.||_2$ stating for the vector norm 2 in $\mathbb{R}^3$.

The motions need to have at least two non-parallel rotation vectors other way the matrix D will remain of rank 2. Indeed if $\vec{l}_A$ is that common rotation vector, then $(\mathsf{D}\,\vec{l}_A = 0)$. Remark that the hand-eye representation in $SE(3)$ does not give any intrinsic constraint for the 3D translation like it gives for the rotation with the trace condition. According to the experiments, the rotation R_X tends to be well estimated, while $\vec{t}_X$ is very noise-sensitive for a small number of motions with little amplitudes.

IV. Classic solving using dual quaternion

A. General description of the dual quaternion formulation

The key idea of this approach is to simultaneously estimate the rotation and the translation. Let consider the Eq.(1) in SE(3). Let $\widehat{\mathbf{a}}$, $\widehat{\mathbf{b}}$ and $\widehat{\mathbf{q}}$ be the unit dual quaternions associated with A, B and X. We can re-write Eq.(1) as

$$\widehat{\mathbf{a}}.\widehat{\mathbf{q}} \;=\; \widehat{\mathbf{q}}.\widehat{\mathbf{b}} \tag{26}$$

Taking into account the multiplication of dual quaternions Eq.(12), and splitting the equation in its real and dual parts, it follows that the real part is:

$$\mathbf{a}.\mathbf{q} \;=\; \mathbf{q}.\mathbf{b} \tag{27}$$

and the dual part is

$$\mathbf{a}'.\mathbf{q} + \mathbf{a}.\mathbf{q}' \;=\; \mathbf{q}.\mathbf{b}' + \mathbf{q}'.\mathbf{b}, \tag{28}$$

Remark that the real part equation is similar to Eq.(19).

B. Specificity of the classic approach

In [6] it is stated that the scalar part of $\widehat{\mathbf{a}}$ is equal to the scalar part of $\widehat{\mathbf{b}}$. According to the screw representation, this means that the motions A and B have the same rotation angle (equality of the real scalar parts) and the same amplitude of pitch (equality of the dual scalar parts). This property is taken into account in the classic solving using dual quaternion to discard redundant equations. Considering the quaternion multiplication of Eq.(4) and the simplification for the redundant parts, the final matrix equation for the classic dual quaternion formulation is obtained as follows :

$$\underbrace{\begin{pmatrix} \tilde{\mathsf{K}}(\mathbf{a}, \mathbf{b}) & 0_{3\times 4} \\ \tilde{\mathsf{K}}(\mathbf{a}', \mathbf{b}') & \tilde{\mathsf{K}}(\mathbf{a}, \mathbf{b}) \end{pmatrix}}_{\mathsf{M}} \begin{pmatrix} \mathbf{q} \\ \mathbf{q}' \end{pmatrix} = 0 \tag{29}$$

With $\tilde{\mathsf{K}}(.,.)$ being defined previously for Eq.(21). The matrix M is 6×8 of rank 5 (c.f. [6]). If we assume N motions, then the global linear matrix is constructed as :

$$\mathsf{P} = \begin{pmatrix} \mathsf{M}_1 \\ \vdots \\ \mathsf{M}_\mathsf{N} \end{pmatrix} \tag{30}$$

P is a $6n \times 8$ matrix that has in general rank 6 in a noise-free case. If all the rotation axis are mutually parallel then the rank falls to 5. In the general case, the algorithm uses a SVD decomposition to deduce the two dual quaternion basis that span the right null space of P. The valid quaternion is computed as the intersection of this null space with the subspace of unit dual quaternions represented by Eq.(5) for the real part and Eq.(14) for the dual part.

C. Solving method for the classic approach

Let denote the two dual quaternions that generates the right null space of P by $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{v}}$. The set of dual quaternion solution of Eq.(30) are described as :

$$\widehat{\mathbf{q}} = \alpha_1\,\widehat{\mathbf{u}} + \alpha_2\,\widehat{\mathbf{v}}, \quad \alpha_1\,\alpha_2 \in \mathbb{R} \tag{31}$$

The two real parameters α_1 and α_2 are sub-optimally determined regarding to the constraints of Eqs. (5, 14). Two second order polynomials $\Gamma_{\widehat{\mathbf{u}},\widehat{\mathbf{v}}}(\lambda)$ and $\Delta_{\widehat{\mathbf{u}},\widehat{\mathbf{v}}}(\lambda)$ are respectively obtained from the unit quaternion condition of Eq.(5) and the orthogonality condition of Eq.(14).

Since α_1 and α_2 never vanish together, we can set $\lambda = \frac{\alpha_1}{\alpha_2}$. The algorithm solves the second order equation $(\Delta_{\widehat{\mathbf{u}},\widehat{\mathbf{v}}}(\lambda) = 0)$ and from the two obtained solutions picks up the one that maximizes the polynomial function $\Gamma_{\widehat{\mathbf{u}},\widehat{\mathbf{v}}}(\lambda)$. Let consider this solution as being λ_0, then α_1 and α_2 are computed as :

$$\alpha_2 = \frac{1}{\Gamma_{\widehat{\mathbf{u}},\widehat{\mathbf{v}}}(\lambda_0)} \;\text{ and }\; \alpha_1 = \lambda_0\,\alpha_2 \tag{32}$$

D. Discussion

By construction and according to the experiments, this algorithm is not very stable to the noise perturbation. Mainly for three reasons :

- It does not use the scalar part in the estimation and in presence of noise the scalar equality is no longer verified.
- Even if the translation is robustly represented through the screw formalism, coupling the estimation makes the rotation part suffer from the noise in the translation estimation. Indeed, in one hand, the rotation is fully characterized by the Eq. (27). In the other hand, in the screw representation the rotation is independent of the translation while the translation depends on the rotation.
- The third reason concerns the constraints for a dual quaternion being unitary. The estimation process suggests to first estimate a two dimensional vector subspace in the dual quaternion space and then compute the intersection of this subspace with the subspace that fulfills the constraint of one-to-one correspondence with $SE(3)$. In a noise-free case or even under small disturbance, this intersection is always one point that corresponds to the solution we are looking for. However, under average noise disturbance there is no guarantee that the intersection is one point. Moreover there is no guarantee that there is an intersection.

V. Our algorithm: Improved Dual Quaternion

A. Specificity of the proposed approach

Given the analysis we drew in the previous sections, the strongest points of our algorithm are the following :

- It uses the dual quaternion formalization of the hand-eye problem that gives a robust representation of the rotation and the translation.
- Since the rotation is independent from the translation in the screw representation and to avoid the influence of the noise of the translation on the rotation we estimate the rotation separately from the translation
- The calibration uses the scalar parts to balance the estimate.
- The one-to-one conditions of correspondence with $SE(3)$, encoded in Eqs.(5) and (14), are optimally taken into account in the estimation.

B. Solving method for the proposed approach

The real part of the dual quaternion formulation of Eq. (27) follows exactly the formulation of Eq. (21), where the scalar parts are included. The dual part of the dual quaternion formulation of Eq. (28) can be written as :

$$\mathsf{K}(\mathbf{a}', \mathbf{b}')\,\mathbf{q} + \mathsf{K}(\mathbf{a}, \mathbf{b})\,\mathbf{q}' = 0 \tag{33}$$

The expression of Eq.(33) contains only the dual part $\mathbf{q}'$ as unknown. Assuming the same set of N motions used to estimate $\mathbf{q}$, we obtain the linear system :

$$\mathsf{L}\,\mathbf{q}' = -\mathsf{L}'\,\mathbf{q} \tag{34}$$

Where L is defined in Eq.(23) and L' is defined as :

$$\mathsf{L}' = \begin{pmatrix} \mathsf{K}(\mathbf{a}'_1, \mathbf{b}'_1) \\ \vdots \\ \mathsf{K}(\mathbf{a}'_N, \mathbf{b}'_N) \end{pmatrix} \tag{35}$$

Moreover, to represent a valid translation in $SE(3)$, $\mathbf{q}'$ has to obey the constraint of Eq.(14). This can be easily formulated as a least squares problem submitted to a linear constraint :

$$\min_{Eq.(14)} ||\mathsf{L}\,\mathbf{q}' + \mathsf{L}'\,\mathbf{q}||_2 \tag{36}$$

With $||.||_2$ being the vector norm 2 in $\mathbb{R}^4$. Such a system can be easily solved using classic methods [8]. The table below summarizes the steps of the Improved Dual Quaternion approach *"I.D.Q"* for hand-eye calibration :

Algorithm 1: *"I.D.Q"* algorithm for hand-eye calibration

Data: N motions $\{(\mathbf{a}_i, \mathbf{a}'_i); (\mathbf{b}_i, \mathbf{b}'_i)\}$, with at least 2 non-parallel rotation axis.

Result: the hand-eye transform.

1 construct the matrix L of Eq.(23), using Eq.(21);

2 compute SVD decomposition of L, and deduce the rotation part $\mathbf{q}$;

3 construct L' using Eq.(35);

4 compute the dual part $\mathbf{q}$ by solving Eq.(36);

5 compute $\vec{t_X}$ using Eq.(16). R_X is directly computed from $\mathbf{q}$.

VI. Synthetic Experiments

This section reports a set of synthetic experiments for assessing the accuracy of different hand-eye calibration approaches. We evaluate the new method presented in section V, against the direct estimation in $SE(3)$ using the formulation of Tsai and Lenz [11], and the original dual quaternion calibration proposed by Daniilidis [6]. The comparison is performed using simulated camera motions around an initial position. If nothing is stated, the motions are randomly generated from a zero mean uniform distribution with a standard deviation of $50\,mm$ in translation and 30 degree in the rotation angle.

The three calibration methods are run over the same sequence of N camera motions, and the estimated hand-eye transform is compared against the ground truth. Each experiment is typically repeated 20 times in order to obtain statistically meaningful results. The graphics plot the root mean square (RMS) error for translation amplitude and rotation angle. The mnemonics "(R,t)", *"C.D.Q"* and *"I.D.Q"* denote respectively the direct estimation in $SE(3)$, the Classical Dual Quaternion algorithm, and the novel approach herein proposed. Since "(R,t)" and *"I.D.Q"* estimate the hand-eye rotation in the same manner, the corresponding curves for the error in θ are overlapped. Please note that the vertical axes of the plots are in base 10 logarithmic scale.

A. Robustness to Noise during Motion Estimation

The motion A of the camera holder (the *"hand"*) is typically determined through encoders and/or optical tracking, while the motion B of the camera (the *"eye"*) is estimated using visual processing. Since visual estimation is noisy and prone to errors, this set of experiment aims comparing the robustness of the different approaches to disturbances in the

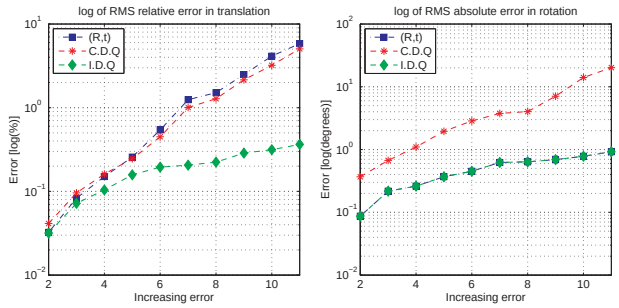


Fig. 2. Accuracy of the hand-eye computation in the presence of noise in the motion estimation. The standard deviation of the additive noise in translation increases from 0 mm to 9 mm in steps of 1 mm, while the noise in rotation increases simultaneously from $0°$ to $2.25°$ in steps of $0.25°$.

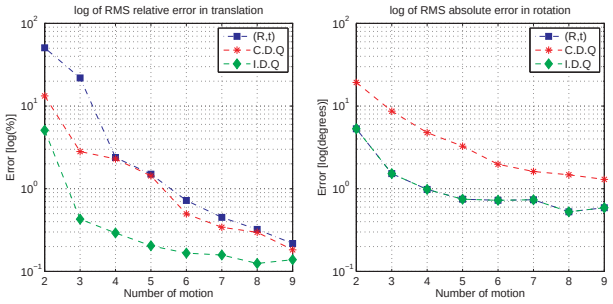


Fig. 3. Accuracy in estimating the hand-eye transformation for an increasing number of input motions

camera motion input. For each run the hand-eye calibration is achieved from a sequence of $N = 9$ motions. The translation and rotation of the camera motion B are disturbed by adding zero mean gaussian noise with increasing standard deviation. Fig. (2) plots the RMS values for the relative error in translation and the absolute error in rotation while estimating the hand-eye transformation. *"I.D.Q"* shows the best behavior followed by *"C.D.Q"* in terms of translation estimate. Therefore, the screw representation is more stable to estimate the translation. The rotation results shows that the separated estimation is more stable regarding to noise disturbance.

B. The effect of the number of motions

The hand-eye calibration for an endoscope is intended to be performed by a non-expert in the operating room, which requires the method to be robust with a minimum number of motions. In this experiment, the number of motion is tuned from $N = 2$ to $N = 9$. The motion range is constant and is parametrized as the previous experiment. For the perturbation, we set $\sigma_{t_\epsilon} = 3\,mm$ and $\sigma_{\theta_\epsilon} = 1.5\,\mathrm{deg}$ which is reasonable in the real experimental conditions. It can be observed in Fig. (3) that all the methods gain in stability when the number of motion increases. However, the proposed method offers better stability for minimal number of motion. Indeed, a high number of motion serves to better stabilize the estimate regarding to the noise. Since our method is robust to noise disturbance it does not need a high N value to give

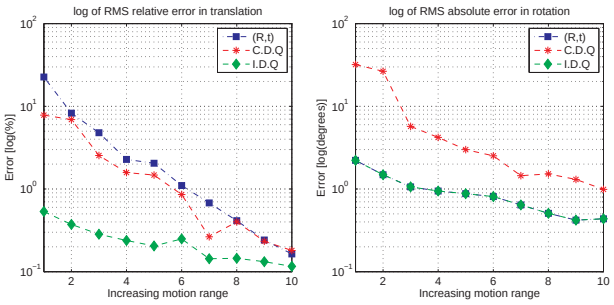


Fig. 4. Effect of the motion amplitude in the estimation error of the hand-eye transformation. The motion amplitude increases in successive trials. The standard deviation of the translation component goes from 10 mm to 100 mm in steps of 10mm, while the angle of the rotation increases from $10°$ to $55°$ in steps of $5°$.

accurate results as the two classic methods do require.

C. The effect of the motion range

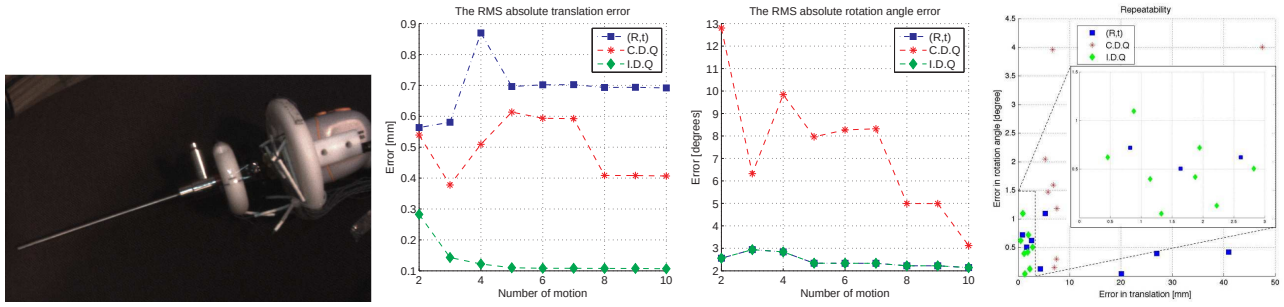
In this experiment we aim analyzing how the amplitude of the motion affects the performance of the three methods. The standard deviations of the uniform distributions from which we randomly pick the input motions is increased in successive trials. The introduced noise is constant and follows the parametrization of the previous experiment. Also, the number of motions is constant and is equal to $N = 5$. Fig. (4) shows that *"I.D.Q"* outperforms the two other methods in terms of accuracy and robustness.Moreover, the (R, t) method suffers of ill-condition matrix under few number of motion or important amount of noise with respect to the range of motion. For the same reasons, the estimation given by the classic dual quaternion is non-stable and can even give imaginary values when solving the equation Eq.(32). Our algorithm bypasses these drawbacks by including the constraints in the linear estimation process.

VII. Real Experimental Results

This algorithm is implemented upon an endoscopic visual framework for knee surgery. The configuration of the endoscope frame (the "hand" pose) is given by the opto-tracker and the camera pose (the "eye" pose) is estimated through the hand-eye transform Fig. (5). Our issue is to estimate the hand-eye transform between those two frames.

In a real experimental context, a direct accurate measurement of the hand-eye transform is not available. Henceforth, we are going to assess the performance of the hand-eye calibration method with two task-dependent evaluations : The first one, also applied in the classical papers [11] and [6], is the ability to predict the camera pose by using only the Opto-Sensor information. The second assessment is the comparison of the repeatability of the calibration result of the three methods.

For both tasks, the camera was moved 25 times with the amounts of translation and rotation varying between 3 and $60\,mm$, and 5 deg and 25 deg respectively. The camera calibration was achieved with *Bouguet Toolbox* [3] that includes pose estimation.



(a) Medical endoscope that is opto-tracked (b) (c) (d)

Fig. 5. A medical endoscopic camera is a conventional CCD camera to which is coupled a special lens system (the rigid scope). The scope is inserted through a small incision into an anatomic cavity (in our case the knee joint), enabling the inspection and visualization of its interior.

In the first experiment, we split the set of images in two subsets for calibration and prediction. We use the $j = 1, \cdots, 11$ first poses to compute the hand-eye transform for an increasing number of corresponding motions $N = 2 \cdots 10$ (two successive poses are equivalent to one motion). Using the hand-eye estimate for each N and the opto-sensor information we predict the camera poses that correspond to the images $j = 12 \cdots 25$ and we compare them to the ones returned by the camera calibration. The RMS absolute translation and rotation errors are reported in Fig. (5-b,c) for the three methods. For a small amplitude and number of motions the results depend on the noise in the camera poses. However, even with this disturbance the results given by the novel method are more accurate than the two other methods starting from a minimum number of $N = 2$ motions.

In the second experiment, we use the whole set of 25 images to obtain a reference estimate of the hand-eye calibration for each of the three methods. Afterward, the set of 24 motions is subdivided in 8 subsets of 3 motions. For each subset and for each method, we estimate the hand-eye transform and compare the result to the corresponding reference estimate. Fig. (5-d) exhibits the absolute rotation error in function of the absolute translation error. It can be observed that *"C.D.Q"* shows good repeatability in translation when compared to "(R,t)". On the other hand, "(R,t)" shows better repeatability in rotation when compared with the *"C.D.Q"*. The proposed approach shows the best repeatability in both rotation and translation.

VIII. Conclusion

The main contribution of this paper is a robust and stable estimation of the hand-eye pose. Our algorithm uses the entire dual quaternion representation. It estimates the rotation separately from the translation and takes advantage of the quaternion representation of the screw to robustly estimates the translation. Moreover, the involved constraints are optimally taken into account during the estimation process. The simulation results show up the accuracy and the stability of our approach regarding to a small number and amplitude of motions. In the context of an endoscopic calibration, the real experimental results emphasize the high accuracy and repeatability when using our algorithm with three motions.

References

[1] P.K. Allen, A. Timcenko, B. Yoshimi, and P. Michelman. Automated tracking and grasping of a moving object with a robotic hand-eye system. *In ICRA*, 9(2):152–165, April 1993.

[2] Joao P. Barreto. Automatic calibration of medical endoscopes using a single image of a planar grid. In *British Machine Vision Conference, BMVC09*, 2009.

[3] J. Y. Bouguet. *Camera Calibration Toolbox for Matlab*. 2003.

[4] Jack C. K. Chou and M. Kamel. Quaternions approach to solve the kinematic equation of rotation of a sensor-mounted robotic manipulator. In *ICRA*, volume 2, pages 656–662, 1988.

[5] Jack C. K. Chou and M. Kamel. Finding the position and orientation of a sensor on a robot manipulator using quaternions. *International Journal of Robotics Research*, 10(2):240–254, June 1991.

[6] Konstantinos Daniilidis. Hand-eye calibration using dual quaternions. *International Journal of Robotics Research*, 18:286–298, 1998.

[7] F. Dornaika and R. Horaud. Simultaneous robot-world and hand-eye calibration. In *ICRA*, volume 14, pages 617–622, 1998.

[8] G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996.

[9] J. B. Kuipers. *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*. Princeton University Press, 2002.

[10] Bruno Siciliano and Oussama Khatib. *Handbook of Robotics*. Springer, 2008.

[11] R.Y. Tsai and R.K. Lenz. Real time versatile robotics hand/eye calibration using 3d machine vision. *ICRA*, 1:554–561, April 1988.

[12] C. Wengert, P. Cattin, J. Duff, and G. Szekely. Markerless endoscopic registration and referencing. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, volume 4190, pages 816–823. Springer, October 2006.

[13] S. Wijesoma, D. Wolfe, and R. Richards. Eye-to-hand coordination for vision-guided robot control applications. *Int. J. Robot. Res.*, 12:65–78, 1993.

[14] A.K.C. Wong, R.V. Mayorga, A. Rong, and X. Liang. A vision based online motion planning of robot manipulators. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 96*, volume 2, pages 940–948 vol.2, Nov 1996.

[15] Shiu Y.C. and Ahmad S. Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form ax=xb. *IEEE Transactions on Robotics and Automation*, 5(1):16–29, 1989.