Laser Steering using Virtual Trifocal Visual Servoing

Nicolas Andreff and Brahim Tamadazte

FEMTO-ST Institute, UMR CNRS 6174 - UFC / ENSMM / UTBM.
Automatic Control and Micro-Mechatronic Systems (AS2M) Department.
24 rue Alain Savary, 25000 Besançon, France

Corresponding author:
firstname.lastname@femto-st.fr

Abstract

This paper focuses on the development of a weakly calibrated three-view based visual servoing control law applied to laser steering process. It proposes to revisit the conventional trifocal constraints governing a three-view geometry for a more suitable use in the design of an efficient trifocal vision-based control. Thereby, an explicit control law is derived, without any matrix inversion, which allows to simply prove the global exponential stability of the control. Moreover, only *"twenty-five lines of code"* are necessary to design a fast trifocal control system. Despite or thanks to the simplicity of the implementation, our control law is fast, accurate, robust to errors on the weak calibration, and exhibits good behavior in terms of convergence and decoupling. This was demonstrated by different experimental validations performed on a test-bench for the steering of laser spot on a 2D and 3D surface using a two degrees-of-freedom commercial piezoelectric mirror, as well as in preliminary cadaver trials using an endoluminal micromirror prototype.

keywords: Multi View Geometry, Trifocal Transfer, Visual Servoing, Laser Surgery, Stability, Robustness.

1 Introduction

Visual servoing techniques, also known as vision-based robot control, use visual information extracted from the images to design a control law (Sanderson and Weiss, 1983; Espiau et al., 1992; Hutchinson et al., 1996). In fact, many vision-based control techniques have emerged with many more realistic applications: smart homes (Liu et al., 2012), mobile

robots (Courbon et al., 2005), industrial manipulators (Lippiello et al., 2007), parallel kinematic mechanisms (Andreff et al., 2007), agricultural machines (Khadraoui et al., 1998), microrobots (Tamadazte et al., 2010; Zhenjiang et al., 2012; Tamadazte et al., 2012), aerospace applications (Coutard et al., 2011), and more recently medical and surgical robotics (Krupa et al., 2009; Rosa et al., 2013; Andreff et al., 2013a).

In the literature, visual servoing approaches (Chaumette and Hutchinson, 2006, 2007) are mainly classified according to: the type of the sensory inputs used to design the control law (e.g., 2D, $2\frac{1}{2}$D, etc.), the configuration of the camera relative to the robot effector (e.g., *eye-in-hand* or *eye-to-hand*), with or without visual tracking algorithm, or quite simply by the number of the vision sensors used in the process.

Several authors (Hespanha et al., 1998; Ruf and Horaud, 1999; Lamiroy et al., 2000; Pari et al., 2010; Alkhalil and Doignon, 2012) have studied stereoscopic visual servoing, using the epipolar geometry between two views. Epipolar geometry or, in case of planar scenes, the homography between two views was also exploited for monocular visual servoing (Rives, 2000; Benhimane and Malis, 2007). However, despite the interest of such approaches, there are still some shortcomings. For instance, the epipolar geometry constraint is difficult to satisfy with short baseline and with planar scenes as shown in (Becerra and Sagüés, 2009; Yang et al., 2014; Montijano et al., 2013). This is the same regarding homography estimation in the case of large baseline in addition to the singularities problems in the system control. To overcome these drawbacks, it is interesting to add an additional sensor (laser, vision system, image, etc.) to build a three-view system. The resulting configuration is commonly called *trifocal geometry* (Hartley and Zisserman, 2006) defined by the trifocal constraint. This geometry is mainly used in 3D scene reconstruction (Liu et al., 2013), 3D pose computation (Becerra and Sagüés, 2009), and also for designing efficient visual servoing control laws (Shademan and Jägersand, 2010; López-Nicolás et al., 2010; Becerra et al., 2011; Sabatta and Siegwart, 2013; Becerra and Sagüés, 2014). It has been proved that the trifocal constraint, contrary to the epipolar constraint, describes the relative geometry of three-view without any knowledge of the observed scene (Hartley and Zisserman, 2006). Note that control system based on trifocal tensors requires, inevitably, substantial matrices manipulation and thus, naturally matrix inversions (often numerically approximated), whose solutions are not obvious. For the moment, this type of control law is commonly used in autonomous personal transportation vehicles and robot manipulators.

Unlike the three-view systems generally studied in the past that consist of three cameras, our approach uses only two cameras along with a laser source. The latter is not used as a depth (Z) estimator, as in (Krupa et al., 2002; Avanzini et al., 2008; Xie et al., 2009) but as a virtual camera (Andreff et al., 2002).

However, in the latter, the laser was fixed with respect to a single camera, thus defining a trivial constant epipolar constraint between the virtual single pixel camera and the real image. In (Andreff et al., 2013a), we addressed the case where the laser has 2 degrees of freedom (pan and tilt) with respect to the real camera, thus scanning the virtual image one pixel at a time. Here, we dig further in multi-view computer vision, since two cameras and a pan-tilt laser are ruled by trifocal geometry, as initially depicted in (Tamadazte and Andreff,

2014). The other contribution of our work resides in the way that the trifocal constraint is taken into account in the design of the visual servoing control law. To do this, we propose to unearth the old trifocal transfer formulation [◦] (Hartley and Zisserman, 2006), which is more adapted form to the design of a control law than the traditional complex mathematical formulations (i.e., 3×3×3 tensor).

In this paper, deepening the analysis proposed in (Tamadazte and Andreff, 2014), we demonstrate that this old formulation allows designing an efficient control law in a very few lines of code (let us say twenty-five, in reference to (Dementhon and Davis, 1995)). This analogy has many advantages:

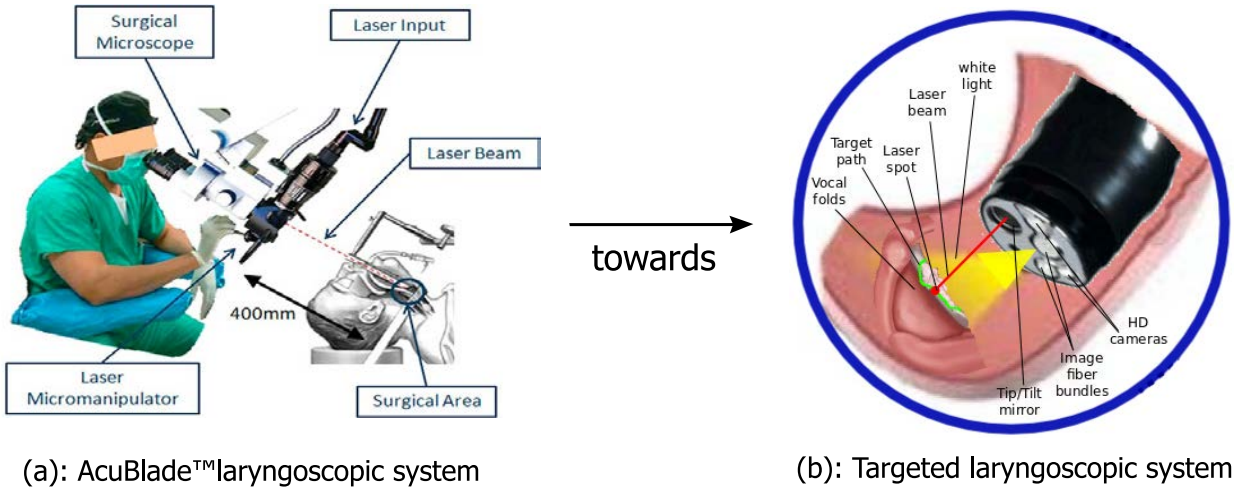
- no matrix inversion nor manipulation of complex mathematical formulas;
- no proprioceptive sensing to know the robot configuration;
- weak calibration, instead of strong camera and hand-eye calibration;
- ability to derive explicit stability conditions.

The specific contributions with respect to (Tamadazte and Andreff, 2014) are more details and discussions, several stability theorems, robustness analysis and extended experiments, including a first in-cadaver trial.

In the remainder of this paper, Section 2 specifies the targeted applications (i.e., medical objectives) of this work. Section 3 reminds the basics of three-view geometry and the derived trifocal constraint and the existing analogy to the tensor formulation. The designed visual servoing control law is detailed in Section 4 while its stability analysis is presented in Section 4.5. Finally, Section 6 shows experimentally the performances i.e., decoupling, convergence, and precision of the developed controller in 2D and 3D targeted scenes.

2 Medical Objectives

The work described in this paper has several applications, beyond laser surgery applications. Vision-based steering can be found an important issue for industrial applications that use laser scanning process. There are several works reported in the literature, especially in vision-based control (often in open-loop mode) supervision of welding process (Sibillano et al., 2009), for example using a robotic arm which embeds the vision system as well as the laser source (Fridenfalk and Bolmsjö, 2004; Huang et al., 2012; Cui et al., 2013). In addition, precise, efficient laser steering can find applications in 2D/3D laser-based micromachining (e.g., microsystems fabrication, semiconductors, glasses, etc) (Rajesh and Bellouard, 2010; Petrak et al., 2011) as well as in free-contact micromanipulation techniques (e.g., laser trapping) (Nahmias and Oddel, 2002; Arai et al., 2004). Otherwise, the developed approach can also address issues in mobile robotics navigation where a laser scanner is associated to a vision system for the 3D perception of the navigation space (Dedieu et al., 2000; Huh et al., 2013). Further applications related to the laser-camera combination emerged in the



(a): AcuBlade™laryngoscopic system

(b): Targeted laryngoscopic system

Figure 1: (a) Current system *versurs* (b) targeted system.

literature, for instance in the control of Unmanned Aircraft Vehicles (UAV) (Wang et al., 2013).

Nevertheless, the objectives of this paper are inscribed in laser surgery in general, and specifically in automatic laser microphonosurgery of the vocal folds. More precisely, it consists of the development of an intuitive endoluminal surgery system for laser ablation and resection of cancerous tissues. Currently, the most successful protocol for vocal folds surgery, widely used in hospitals is certainly the suspension laryngoscopic technique which consists of a straight-rigid laryngoscope, a stereomicroscope, a set of specific and miniatures surgical instruments, a laser source, and a foot-pedal controller device to activate the laser (Jako, 1970; Eckel et al., 2003). These systems allow performing more precise diagnostics and microsurgeries despite many problems for patients and surgeons. For instance, the AcuBlade™system (which claims to be a robotic system) uses a stereomicroscope and a laser source placed outside the mouth, at 400 mm from the vocal cords (Fig. 1(a)). This disposition involves the projection, in straight line, of the laser beam on the soft tissue through the patient's mouth, the laryngoscope tunnel and the larynx itself. This causes a very uncomfortable position for the patient with an extreme extension of the neck, which makes it painful several days after the operation (Aloy and Grasl, 2013). It also raises safety concerns, related to bad alignment of the laser and the throat (burning risks). Apart from this, the requirement of dexterity and considerable expertise of the surgeon to perform an intervention of the vocal cords is highly required. It is true, among other, because the laser displacements are performed in an open-loop control scheme.

To overcome these numerous drawbacks, the μRALP project[1] proposes a new concept of a more intuitive and adjustable laryngoscope. This consists of an endoscope which is equipped with a two degrees-of-freedom (dof) piezoelectric mirror, a fixed mirror, two image bundles, two high-definition RGB cameras (for the surgeon visualization), a light source, and a laser

[1] www.microralp.eu

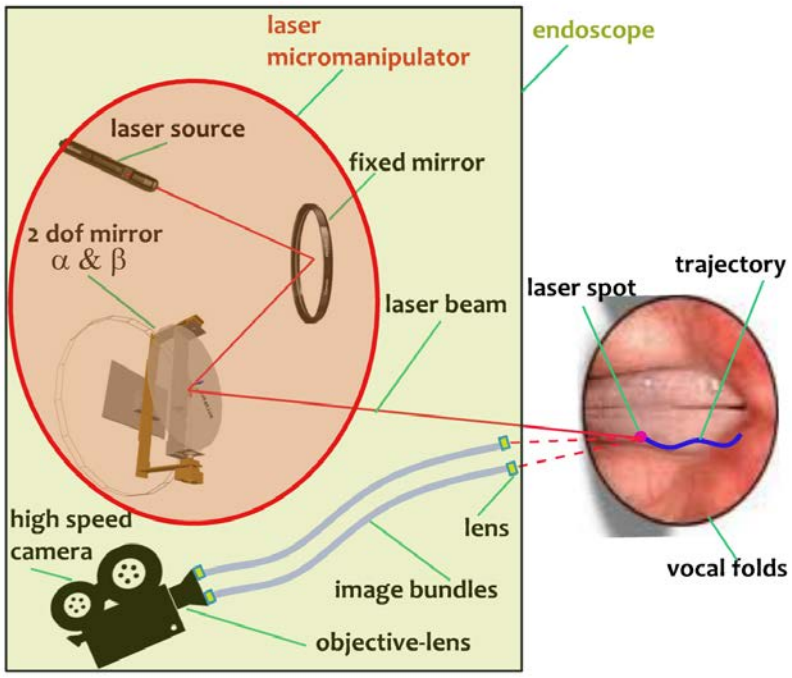


Figure 2: Schematic view of the laser steering system with two cameras.

source (Fig. 1(b)). This new flexible laryngoscope system will be placed directly inside the larynx close (i.e., 20 mm) to the vocal folds. In the same way, the laser spot displacements will no longer be controlled in semi-automatic nor in open-loop control strategy as are currently performed. The endoscope is especially designed to ensure a stereoscopic visualization of the vocal cords to which the laser source is added (co-axial visible laser and CO2 surgical laser) as a virtual camera to form a three-view system. This specific configuration allows a laser visual servoing system which will ensure precise, robust and accurate execution of laser excision paths and ablation patterns defined intraoperatively by the surgeon through an intuitive Surgeon-Robot Interface, such as the one proposed in (Dagnino et al., 2012). This offers an automatic correction of aiming errors resulting from intrinsic mechanical characteristics of the laser microscanner, piezoelectric actuators drift, or calibration inaccuracies. Moreover, the tissue deformation during surgery intervention can be tracked in real-time in order to provide an update to the reference trajectory (Schoob et al., 2013). Also, the use of the visual feedback allows to alleviate the mechanical design of the mirror mechanism from the need of joint encoders (Shoham and Koren, 1988; Marchand et al., 2002; Andreff and Martinet, 2009), which are a remarkable feature knowing the size constraints. Therefore, the automatic steering of the laser displacements will significantly increase the adaptability, accuracy and safety of laser phonomicrosurgeries compared to the currently existing technologies.

For a better understanding the functioning of the laryngoscopic system mentioned in this work, as well as the placement of the various devices that compose it, one can refer to Fig. 2 which shows a schematic view of the hardware. In the final device, the two cameras will be replaced by fiber optic bundles (50K fibers for each) which provide two stereoscopic images

of the vocal folds onto a high-speed (10 000 frames per second) CMOS camera. Indeed, the laser surgery operates by bringing a high energy density onto the cells. If the laser sweeps the surface fast enough, the amount of energy is just enough to sublimate (vaporize) the cells; if not, the energy transfers into heat in the surrounding tissue and the latter carbonizes (Shimokita et al., 2008). As a consequence, laser surgery implies high bandwidth (at least 200 Hz) sensing devices and control laws, namely high-frequency visual servoing which explains the use of the high-speed camera.

3 Trifocal Geometry

The trifocal geometry is defined by a 3×3×3 array tensor which includes all projective geometric relationships among three-view of the three cameras. For instance, it allows to link the coordinates of corresponding points $\boldsymbol{p}_i = (x_i, y_i)^\top$ or lines (l_i) in the other views regardless of the scene structure but only in the relative poses of the cameras. The trifocal tensor can be *considered* as the generalization of the fundamental matrix in a three-view system.

3.1 Known Notations

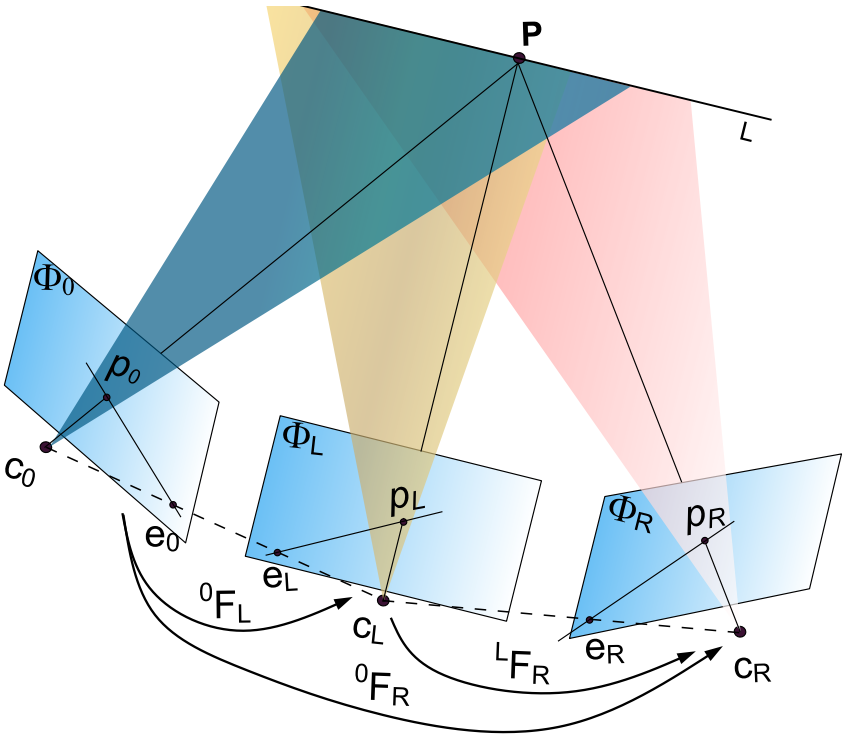


Figure 3: Trifocal geometry principle.

Let three cameras with optical centers $\boldsymbol{c}_0$, $\boldsymbol{c}_L$ and $\boldsymbol{c}_R$ observe a 3D point ${}^0\mathbf{P} = (X, Y, Z)^\top$

which is projected in 2D points $\boldsymbol{p}_0 = (x, y)^\top$, $\boldsymbol{p}_L = (x_L, y_L)^\top$, and $\boldsymbol{p}_R = (x_R, y_R)^\top$ in the images planes Φ_0, Φ_L, and Φ_R, respectively. The line segments $(\boldsymbol{c}_0\boldsymbol{c}_L)$, and $(\boldsymbol{c}_L\boldsymbol{c}_R)$ define the baselines between each camera pair. The intersections of these baselines with the image planes give the epipolar points $\boldsymbol{e}_0$, $\boldsymbol{e}_L$, and $\boldsymbol{e}_R$. Also, $(\boldsymbol{e}_0\boldsymbol{p}_0)$, $(\boldsymbol{e}_L\boldsymbol{p}_L)$, and $(\boldsymbol{e}_R\boldsymbol{p}_R)$ are called epipolar lines (Fig. 3). There are mathematical relations between the epipolar lines $(\boldsymbol{e}_0\boldsymbol{p}_0)$, $(\boldsymbol{e}_L\boldsymbol{p}_L)$, and $(\boldsymbol{e}_R\boldsymbol{p}_R)$ and the 2D point $\boldsymbol{p}_0$, commonly called epipolar constraints which are given by

$$\tilde{\boldsymbol{p}}_0^\top {}^0\mathbf{F}_L \tilde{\boldsymbol{p}}_L = 0 \tag{1}$$

$$\tilde{\boldsymbol{p}}_L^\top {}^L\mathbf{F}_R \tilde{\boldsymbol{p}}_R = 0 \tag{2}$$

$$\tilde{\boldsymbol{p}}_0 {}^0\mathbf{F}_R \tilde{\boldsymbol{p}}_R^\top = 0 \tag{3}$$

where, ${}^i\mathbf{F}_j$ is the fundamental matrix between camera i and camera j and $\tilde{\boldsymbol{p}}_i$ the homogenous representation of $\boldsymbol{p}_i$ ($\tilde{\boldsymbol{p}}_i = (\boldsymbol{p}_i, 1)^\top$).

It is also possible to represent these constraints by the trifocal tensor of a point-point-point transformation defined in (Hartley and Zisserman, 2006) by

$$[\tilde{\boldsymbol{p}}_L]_\times \left(\sum_{i=1}^{3} \tilde{\boldsymbol{p}}_i \mathbf{T}_i \right) [\tilde{\boldsymbol{p}}_R]_\times = 0_{3\times 3} \tag{4}$$

where $[\boldsymbol{v}]_\wedge$ is the skew-symmetric matrix associated to the vector cross-product ($\times$) by $\boldsymbol{v}$ and where the set of 3×3 matrices $\mathbf{T}_1$, $\mathbf{T}_2$, and $\mathbf{T}_3$ represent the trifocal tensor $\mathcal{T}_{3\times 3\times 3}$. Therefore, it is possible to write the point-point-point transformation as

$$[\tilde{\boldsymbol{p}}_L]_\times (\tilde{\boldsymbol{p}}_i \mathcal{T}_i^{jk}) [\tilde{\boldsymbol{p}}_R]_\times = 0_{3\times 3} \tag{5}$$

with i,j,$k \in [1;3]$.

Although mathematically sound and complete, this formulation is far from being water-clear for a roboticist.

3.2 Another formulation of the Trifocal Geometry

As shown in Fig. 4, it is possible to define an analogy between our system and the conventional three-view system illustrated in Fig. 3 where the third view is a virtual image made by the pan-tilt laser. The approach shown here is based on a *vectorial writing* of the visual servoing problem using the trifocal constraint. This is mainly to avoid the conventional equation (4) involving a complex mathematical formulation based on tensors. Therefore, we demonstrate that the use of the older expression of the trifocal transfer (Hartley and Zisserman, 2006, p. 380) is better suited for the design of the visual servoing control law, since it allows to avoid numerical matrix inversions at all. Let us call ${}^0\underline{\boldsymbol{z}}$ a unit-vector representing the direction of the laser beam/line from the steering mirror towards the target (e.g., vocal folds), ${}^0\mathbf{P}$ the 3D position of the spot laser in the vocal folds, and $\tilde{\boldsymbol{p}}$ its projection in the image. They are related by

$${}^0\mathbf{P} = d\,{}^0\underline{\boldsymbol{z}} = Z\tilde{\boldsymbol{p}}_0 \tag{6}$$

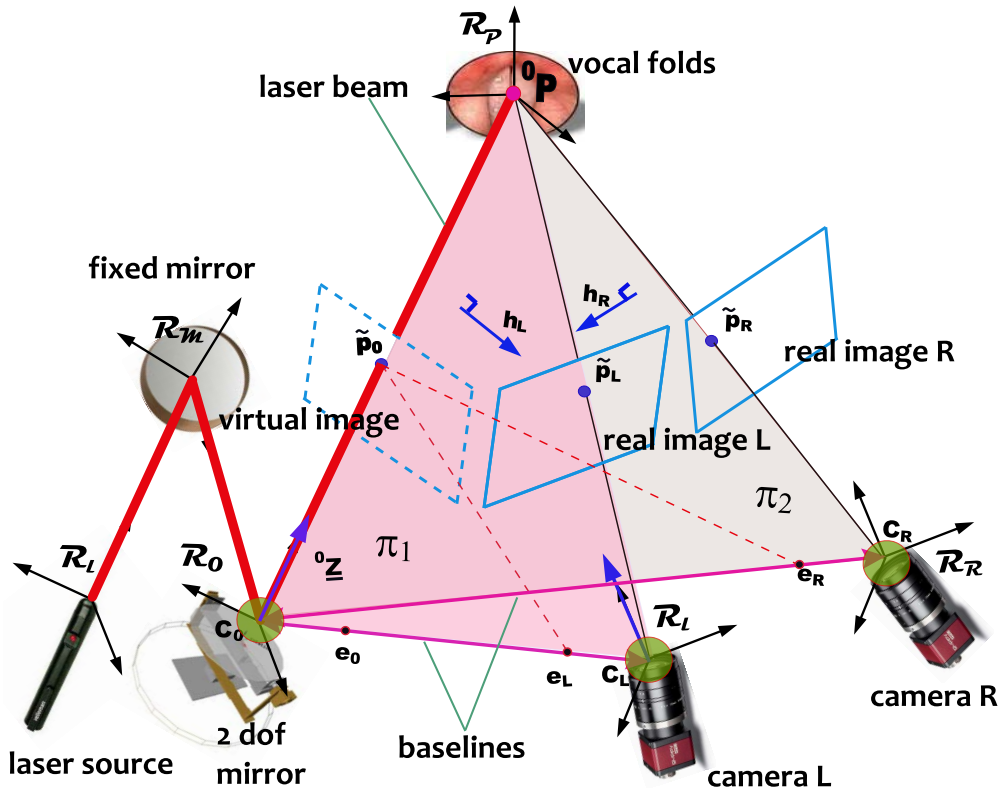


Figure 4: Analogy with trifocal geometry.

with d being the distance between the mobile mirror and the targeted surface, $\tilde{\boldsymbol{p}}$ is the projection of ${}^0\mathbf{P}$ on the virtual image and Z is the depth (i.e., the distance parallel to the virtual optical axis). It has to be highlighted that d is not easy to determine unlike the vector ${}^0\underline{\boldsymbol{z}}$ which is obtainable from the stereoscopic system, as shown later. Therefore, we will not need d nor Z any more in this paper.

From (6), it is easy to see that ${}^0\underline{\boldsymbol{z}}$ and $\tilde{\boldsymbol{p}}_0$ are projectively similar.

$$ {}^0\underline{\boldsymbol{z}} \equiv \tilde{\boldsymbol{p}}_0 \tag{7}$$

Actually, $\tilde{\boldsymbol{p}}_0$ is the perspective projection on a plane, while ${}^0\underline{\boldsymbol{z}}$ is the perspective projection on a sphere. From the layout shown in Fig. 4, it is possible to rewrite the trifocal constraint (1), (2), (3) and (7) under the following form

$$ \tilde{\boldsymbol{p}}_L{}^\top \, {}^L\mathbf{F}_R \, \tilde{\boldsymbol{p}}_R = 0 \tag{8}$$

$$ {}^0\underline{\boldsymbol{z}}^\top \, {}^0\mathbf{F}_R \, \tilde{\boldsymbol{p}}_R = 0 \tag{9}$$

$$ {}^0\underline{\boldsymbol{z}}^\top \, {}^0\mathbf{F}_L \, \tilde{\boldsymbol{p}}_L = 0 \tag{10}$$

Leaving aside (8), let us introduce the notations $\boldsymbol{h}_R$ and $\boldsymbol{h}_L$ as

$$ \boldsymbol{h}_R = {}^0\mathbf{F}_R \tilde{\boldsymbol{p}}_R \qquad \boldsymbol{h}_L = {}^0\mathbf{F}_L \tilde{\boldsymbol{p}}_L \tag{11}$$

which encode the coefficients of the epipolar lines $(\boldsymbol{e}_R\tilde{\boldsymbol{p}}_0)$ and $(\boldsymbol{e}_L\tilde{\boldsymbol{p}}_0)$.

Consequently, from (9) and (10), it is trivial to see that the cross-product $\boldsymbol{h}_R \times \boldsymbol{h}_L$ is parallel to ${}^0\underline{\boldsymbol{z}}$ (i.e., ${}^0\underline{\boldsymbol{z}} \sim \boldsymbol{h}_R \times \boldsymbol{h}_L$). To avoid scale issues, this can be formulated thanks to the cross-product

$$ {}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) = 0 \tag{12} $$

Thereby, the point-point-point transformation formulation (trifocal tensor) given in (5) can be replaced by

$$ {}^0\underline{\boldsymbol{z}} \times \left({}^0\mathbf{F}_R\tilde{\boldsymbol{p}}_R \times {}^0\mathbf{F}_L\tilde{\boldsymbol{p}}_L \right) = 0 \tag{13} $$

This formula represents the trifocal constraint expressed in an algebraic form using vector products of ${}^0\underline{\boldsymbol{z}}$, $\tilde{\boldsymbol{p}}_L$ and $\tilde{\boldsymbol{p}}_R$, contrarily to the matrix formulation of the trifocal constraint given in (4). Therefore, it appears that the new expression of trifocal constraint allows for simple manipulation and time-derivation in order to design the new controller.

This formulation has the drawback, with respect to the trifocal tensor, that it does not allow to build ${}^0\underline{\boldsymbol{z}}$ from $\tilde{\boldsymbol{p}}_L$ and $\tilde{\boldsymbol{p}}_R$ when ${}^0\underline{\boldsymbol{z}}$ is coplanar with $\boldsymbol{e}_L$ and $\boldsymbol{e}_R$ (*i.e.* when ${}^0\underline{\boldsymbol{z}}$ crosses the line drawn between those epipoles in the virtual image.

3.3 Analysis

The epipolar lines $\boldsymbol{h}_L$ and $\boldsymbol{h}_R$ can be computed either using the desired laser spot in both images and the laser-image fundamental matrices, as above (11) or using the laser-image epipoles and the desired laser beam direction (Fig. 6). To do so, remember that the coefficients of an image line is obtained by the cross-product between two points on the line

$$ \boldsymbol{h}_L = \boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}} \qquad \boldsymbol{h}_R = \boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}} \tag{14} $$

where ${}^0\mathbf{F}_L\boldsymbol{e}_L = 0$, ${}^0\mathbf{F}_R\boldsymbol{e}_R = 0$ and the sign of the fundamental matrices and epipoles must match i.e., be chosen such that

$$ ({}^0\mathbf{F}_L\tilde{\boldsymbol{p}}_L)^\top(\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}}) > 0 \qquad ({}^0\mathbf{F}_R\tilde{\boldsymbol{p}}_R)^\top(\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}}) > 0 \tag{15} $$

so that (11) and (14) are equivalent up to a *strictly positive* factor. From (14), we get trivially

$$ \boldsymbol{h}_R \times \boldsymbol{h}_L = (\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}}) \times (\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}}) \tag{16} $$

Now develop this expression, using the well-known formula for the vector triple product

$$ \boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}\left(\boldsymbol{a}^\top\boldsymbol{c}\right) - \boldsymbol{c}\left(\boldsymbol{a}^\top\boldsymbol{b}\right) \tag{17} $$

and obtain, after some development and rearrangement,

$$ \boldsymbol{h}_R \times \boldsymbol{h}_L = -\,{}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)\,{}^0\underline{\boldsymbol{z}} \tag{18} $$

In that expression appears the signed distance ${}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)$ between ${}^0\underline{\boldsymbol{z}}$ to the line passing by the two epipolar points (Fig. 5). Thereby,

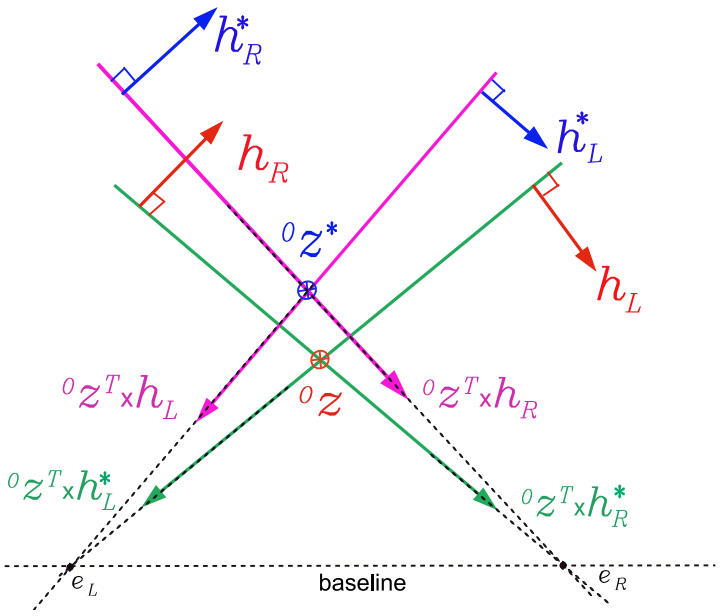


Figure 5: Sign issues related to which side of the line joining the two virtual epipoles the laser beam lies.

- this explains that the singularity for the trifocal constraint occurs when the laser beam crosses the line joining the two cameras centers of projection, since then $\boldsymbol{h}_R \times \boldsymbol{h}_L$ becomes $\mathbf{0}$.
- Outside this singularity, the following relations hold (Fig. 6)

$$\|\boldsymbol{h}_R \times \boldsymbol{h}_L\| = |{}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)| \tag{19}$$

$${}^0\underline{\boldsymbol{z}} = \epsilon \frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} \tag{20}$$

with ϵ = - sign$({}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R))$.

4 Visual Servoing

Now, we have all the elements for designing the new virtual trifocal-based control law using the proposed analogy (12).

4.1 Kinematic Constraints

Firstly, the time derivative of the old trifocal constraint (12) yields

$${}^0\dot{\underline{\boldsymbol{z}}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) + {}^0\underline{\boldsymbol{z}} \times \left(\dot{\boldsymbol{h}}_R \times \boldsymbol{h}_L\right) + {}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L\right) = 0 \tag{21}$$

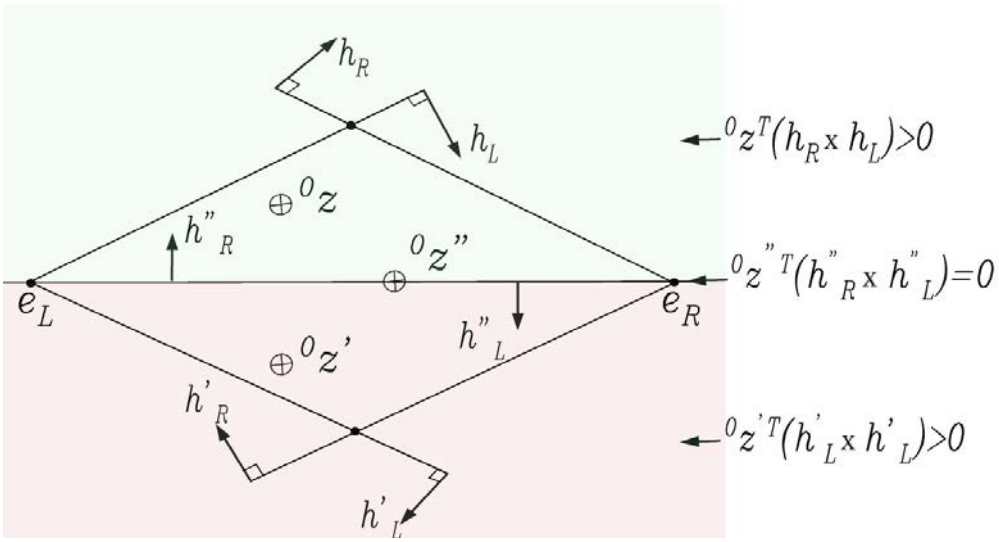


Figure 6: Different cases of the location of ${}^0\underline{\boldsymbol{z}}$ regarding the baseline $(\boldsymbol{e}_L \boldsymbol{e}_R)$.

Since the trifocal geometry is fixed, we infer that $\dot{\boldsymbol{h}}_R = {}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R$ and $\dot{\boldsymbol{h}}_L = {}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L$, which turn the latter equation into

$$ {}^0\dot{\underline{\boldsymbol{z}}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) + {}^0\underline{\boldsymbol{z}} \times \left(({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) \times \boldsymbol{h}_L \right) + {}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) = 0 \tag{22} $$

or, equally, into

$$ {}^0\dot{\underline{\boldsymbol{z}}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) = - {}^0\underline{\boldsymbol{z}} \times \left(({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) \times \boldsymbol{h}_L \right) - {}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \tag{23} $$

which can be reorganized as follows

$$ {}^0\dot{\underline{\boldsymbol{z}}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) = {}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \tag{24} $$

First, dividing each side of the equation (24) by $\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|$ yields

$$ {}^0\dot{\underline{\boldsymbol{z}}} \times \frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} = \frac{{}^0\underline{\boldsymbol{z}}}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \tag{25} $$

Second, using (20) on both sides and inverting the cross-product ${}^0\dot{\underline{\boldsymbol{z}}} \times {}^0\underline{\boldsymbol{z}}$, we finally obtain a useful kinematic relationship between the velocities $\dot{\tilde{\boldsymbol{p}}}_L$ and $\dot{\tilde{\boldsymbol{p}}}_R$ of the laser spot in each real image $\mathbf{I}_L$ and $\mathbf{I}_R$, respectively, and the laser beam change of orientation ${}^0\dot{\underline{\boldsymbol{z}}}$

$$ {}^0\underline{\boldsymbol{z}} \times {}^0\dot{\underline{\boldsymbol{z}}} = - \frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|^2} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \tag{26} $$

4.2 Control Law

Now, it is necessary to link the laser beam velocities ${}^0\underline{\dot{\boldsymbol{z}}}$ to the rotation velocity of the pan-tilt steering mirror denoted by $\boldsymbol{\omega}$.

Knowing that ${}^0\underline{\dot{\boldsymbol{z}}}$ is a unit vector (*i.e.*, ${}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}} = 1$ and ${}^0\underline{\dot{\boldsymbol{z}}}^\top {}^0\underline{\boldsymbol{z}} = 0$), we can write

$$ {}^0\underline{\dot{\boldsymbol{z}}} = \boldsymbol{\omega} \times {}^0\underline{\boldsymbol{z}} \tag{27} $$

the analytic solution of which is

$$ \boldsymbol{\omega} = {}^0\underline{\boldsymbol{z}} \times {}^0\underline{\dot{\boldsymbol{z}}} + k\,{}^0\underline{\boldsymbol{z}}, k \in \Re \tag{28} $$

However, a rotation of the laser beam around its axis is not significant. Therefore, k is chosen equal to 0, and thus

$$ \boldsymbol{\omega} = {}^0\underline{\boldsymbol{z}} \times {}^0\underline{\dot{\boldsymbol{z}}} \tag{29} $$

Therefore, by identifying (26) and (29), we can easily deduce the expression of the mirror velocity as a function of the laser spot velocities

$$ \boldsymbol{\omega} = -\frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|^2} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \tag{30} $$

For an exponential decay of the cost-function $\mathbf{e} = \begin{pmatrix} \tilde{\boldsymbol{p}}_L - \tilde{\boldsymbol{p}}_L^* \\ \tilde{\boldsymbol{p}}_R - \tilde{\boldsymbol{p}}_R^* \end{pmatrix}$, we introduce a first-order behavior of the error, in both images, between the current position $\tilde{\boldsymbol{p}}$ and the desired position $\tilde{\boldsymbol{p}}^*$ of the laser spot

$$ \dot{\tilde{\boldsymbol{p}}}_{\boldsymbol{i}} = -\lambda(\tilde{\boldsymbol{p}}_{\boldsymbol{i}} - \tilde{\boldsymbol{p}}_i^*) + \dot{\tilde{\boldsymbol{p}}}_i^*, i \in \{L, R\} \tag{31} $$

where λ is a positive gain and $\dot{\tilde{\boldsymbol{p}}}_i^*$ is the feed-forward term in case of trajectory tracking which is directly given by the surgeon through the Robot-Surgeon Interface.

Using (31), (30) can be rewritten as

$$ \begin{aligned} \boldsymbol{\omega} \;=\; & \lambda \frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|^2} \times \left(\boldsymbol{h}_L \times \left({}^0\mathbf{F}_R (\tilde{\boldsymbol{p}}_{\boldsymbol{r}} - \tilde{\boldsymbol{p}}_R^*) \right) - \boldsymbol{h}_R \times \left({}^0\mathbf{F}_L (\tilde{\boldsymbol{p}}_{\boldsymbol{l}} - \tilde{\boldsymbol{p}}_L^*) \right) \right) \\ & - \frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|^2} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R^*) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L^*) \right) \end{aligned} \tag{32} $$

Reminding that $\boldsymbol{h}_i = {}^0\mathbf{F}_i \tilde{\boldsymbol{p}}_{\boldsymbol{i}}$, and introducing the notations $\boldsymbol{h}_i^* = {}^0\mathbf{F}_i \tilde{\boldsymbol{p}}_i^*$ and $\dot{\boldsymbol{h}}_i^* = {}^0\mathbf{F}_i \dot{\tilde{\boldsymbol{p}}}_i^*$, equation (32) simplifies into the final control law

$$ \boldsymbol{\omega} = \underbrace{-\lambda \frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|^2} \times \left(\boldsymbol{h}_L \times \boldsymbol{h}_R^* - \boldsymbol{h}_R \times \boldsymbol{h}_L^* \right)}_{first\ order\ behavior\ of\ the\ error} \underbrace{- \frac{\boldsymbol{h}_R \times \boldsymbol{h}_L}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|^2} \times \left(\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^* - \boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^* \right)}_{feed-forward\ term} \tag{33} $$

Note that this control is fully based on image measurements. Note also that this control is expected to decouple the convergence of each component of the laser spot position in

the images. Namely each of the components should converge with the same behaviour – hopefully, an exponential decay in the image since we have imposed first-order dynamics in (31) – at the same rate – since we have chosen a unique control gain. Such a decoupling appears clearly in visual servoing control laws thanks to the numerical (pseudo-)inversion of the interaction matrix. It is less obvious here because we inverted the motion equation (that is, the underlying interaction matrix) at hand. However, since we have not made any simplifying assumption in this inversion, it is an exact inverse and thus the system is expected to be decoupled.

4.3 Robot kinematics

Finally, $\boldsymbol{\omega}$ is converted to joint velocities $\dot{q}$ through the inverse differential kinematic matrix $\mathbf{D}^{-1}$ inverse robot Jacobian of the mobile mirror mechanism as $\boldsymbol{\omega} = \mathbf{D}^{-1}\dot{\mathbf{q}}$, where $\mathbf{D}^{-1}$ depends on how the pan-tilt motion is technically made. Note that in the case where the mirror mechanism has parallel kinematics, $\mathbf{D}^{-1}$ has two useful features (Paccot et al., 2009):

- it has an algebraic expression $\mathbf{D}^{-1} = \mathbf{D}^{inv}$ (no matrix inversion);
- it depends on the Cartesian configuration $\mathbf{D}^{-1} = \mathbf{D}^{inv}({}^0\underline{\boldsymbol{z}})$.

4.4 Implementation

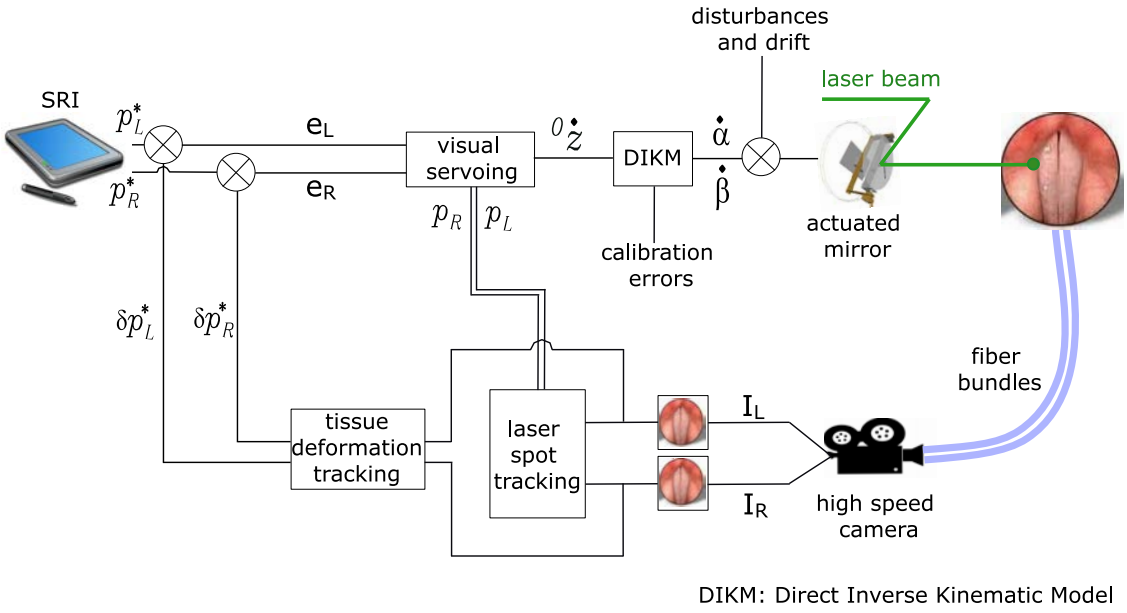


Figure 7: Block diagram showing the operation of the phonomicrosurgical system.

Concerning the implementation, the proposed controller can be summarized by Fig. 7. The "tissue deformation tracking" block estimates the time-variation (e.g., $\dot{\boldsymbol{p}}_i^*$) of the tissue

during the intervention. These deformations can be computed through a model-based implementation using thin plate splines (TPS) and fast optimization implemented in a GPU (Graphics Processing Unit) developed in the μRALP framework by (Schoob et al., 2014). The "disturbance and drift" input concerns especially the nonlinearities (hysteresis, drift, etc.) of the piezoelectric actuators which equip the mobile mirror. These are compensated for by the real-time vision-feedback control.

There is given the source code of the proposed controller. As can be seen the implementation of the control law takes only a few tens lines of code (including comments and software quality) as promised.

```
#include <visp/vpColVector.h>
vpColVector computeTrifocalLaw(const vpColVector& ImagePoints[2],
                const vpColVector& ImagePoints_d[2],
                const vpFundamentalMatrix& F[2])
{
  // Compute the control law
  vpColVector error[2];
  vpColVector ImagePointsdot[2];

  // 1) Assume an exponential decay in both images
  for (unsigned int img=0; img<2; img++) {
    error[img] = (ImagePoints[img] - ImagePoints_d[img]);
    ImagePointsdot[img] = -1*lambda*error[img];
 }

  // 2) Convert the image velocity in the mirror velocity => computes udot

  // Consider the trifocal tensor constraint:
  //   u x { (F[0]*ImagePoints[0]) x (F[1]*ImagePoints[1]) } = 0
  // or
  //   u x ( h[0] x h[1] ) = 0
  // where h[img] is the epipolar line associated to
  //       ImagePoints[img] in the "virtual" camera frame
  //       associated to the mirror

  vpColVector  h[2];
  for (unsigned int img=0; img<2; img++) {
    h[img]=F[img]*ImagePoints[img];
  }

  // Differentiate the constraint with respect to time:
  //  udot x (h[0] x h[1]) + u x (hdot[0] x h[1])
  //  + u x (h[0] x hdot[1])  = 0
  // which is equivalent to:
  //  udot x (h[0] x h[1])  =  u x (h[1] x hdot[0]) - u x (h[0] x hdot[1])
  // or,
  //  (h[0] x h[1]) x udot  = - u x (h[1] x hdot[0]) + u x (h[0] x hdot[1])

  vpColVector hdot[2];
  for (unsigned int img=0; img<2; img++) {
```

```
    hdot[img]= F[img] * ImagePointsdot[img];
  }

  vpColVector h0xh1xudot;
  h0xh1xudot = - vpColVector::cross(u, vpColVector::cross(h[1],hdot[0]))
               + vpColVector::cross(u, vpColVector::cross(h[0],hdot[1]));

  // 3) Compute the rotation velocity of the mirror
  // udot = Omega x u (because u^T u = 1)
  // So, Omega = u x udot
  // but (trifocal constraint)
  //   u ~ h0xh1  and  ||u|| = 1
  // thus, Omega =  (h0xh1xudot) / norm(h0xh1)
  //
  // To fix the sign, we must take the sign of u^T h0xh1

  vpColVector h0xh1;
  h0xh1 = vpColVector::cross(h[0],h[1]);

  if ( (h0xh1.sumSquare() /
   (h[0].sumSquare() * h[1].sumSquare() ))    < 0.01 ) {
    std::cerr << "Singularity!" << std::endl;
    throw;
  }

  vpColVector Omega(3);
  try {
    Omega = h0xh1xudot / sqrt(h0xh1.sumSquare());
  } catch(...) {
    throw;
  }
  return Omega;
}
```

4.5 Stability Analysis

Proposition 1 (Stability condition of the point to point response) *The control in (33) is exponentially stable to a step response ($\forall t > 0,\ {}^0\dot{\underline{\boldsymbol{z}}}^*(t) = 0$) if the laser does not cross the line passing through the two laser-image epipoles.*

Proof:

Let us consider the following Lyapunov function candidate

$$V = \frac{1}{2}\| {}^0\underline{\boldsymbol{z}} - {}^0\underline{\boldsymbol{z}}^*\|^2 \tag{34}$$

where ${}^0\underline{\boldsymbol{z}}^*$ is the desired laser direction which can be obtained using (20)

$${}^0\underline{\boldsymbol{z}}^* = \epsilon^* \frac{\boldsymbol{h}_R^* \times \boldsymbol{h}_L^*}{\|\boldsymbol{h}_R^* \times \boldsymbol{h}_L^*\|} \tag{35}$$

The Lyapunov function candidate is trivially strictly positive when away from the desired orientation and 0 only in the latter. We now need to prove that its derivative is negative (*asymptotic stability*) and bounded by an exponentially decreasing function (*exponential stability*).

The time derivative of V is

$$\dot{V} = ({}^0\underline{\boldsymbol{z}} - {}^0\underline{\boldsymbol{z}}^*)^\top({}^0\dot{\underline{\boldsymbol{z}}} - {}^0\dot{\underline{\boldsymbol{z}}}^*) \tag{36}$$

which develops as

$$\dot{V} = \underbrace{{}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}}_{=0} - \underbrace{{}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^*}_{=0} - {}^0\underline{\boldsymbol{z}}^{*\top} {}^0\dot{\underline{\boldsymbol{z}}} + \underbrace{{}^0\dot{\underline{\boldsymbol{z}}}^{*\top} {}^0\dot{\underline{\boldsymbol{z}}}^*}_{=0} \tag{37}$$

where the above terms cancel either because the desired orientation is, here, assumed constant or because the derivative of a unit vector is orthogonal to the latter. Consequently,

$$\dot{V} = -{}^0\underline{\boldsymbol{z}}^{*\top} {}^0\dot{\underline{\boldsymbol{z}}} \tag{38}$$

Inserting (27) into the latter and reordering the scalar triple product (or mixed product) yields

$$\dot{V} = -({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \boldsymbol{\omega} \tag{39}$$

Now, replace $\boldsymbol{\omega}$ by the control (33) with a constant desired laser position and use (20)

$$\dot{V} = \frac{\lambda\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big({}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \boldsymbol{h}_R^* - \boldsymbol{h}_R \times \boldsymbol{h}_L^*) \Big) \tag{40}$$

Let us now develop the $\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \boldsymbol{h}_R^*)$ term. To do so, we use the well-known formula for the vector triple product (17). If we consider that $\boldsymbol{a} = {}^0\underline{\boldsymbol{z}}$, $\boldsymbol{b} = \boldsymbol{h}_L$, and $\boldsymbol{c} = \boldsymbol{h}_R^*$, then

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \boldsymbol{h}_R^*) = \boldsymbol{h}_L({}^0\underline{\boldsymbol{z}}^\top \boldsymbol{h}_R^*) - \boldsymbol{h}_R^*({}^0\underline{\boldsymbol{z}}^\top \boldsymbol{h}_L) \tag{41}$$

Remember that ${}^0\underline{\boldsymbol{z}}$ is orthogonal to $\boldsymbol{h}_L$ and the latter becomes

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \boldsymbol{h}_R^*) = \boldsymbol{h}_L \, {}^0\underline{\boldsymbol{z}}^\top \boldsymbol{h}_R^* \tag{42}$$

which can be developed using (14)

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \boldsymbol{h}_R^*) = (\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}}) \, {}^0\underline{\boldsymbol{z}}^\top (\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}}^*) \tag{43}$$

and reordered as

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \boldsymbol{h}_R^*) = -(\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}}) \boldsymbol{e}_R^\top ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) \tag{44}$$

Similarly, we have ${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L^*) = -(\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}}) \boldsymbol{e}_L^\top ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)$ and, thus, (40) becomes

$$\dot{V} = \frac{\lambda\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big(-(\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}}) \boldsymbol{e}_R^\top ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) + (\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}}) \boldsymbol{e}_L^\top ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) \Big) \tag{45}$$

Factorizing ${}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*$ on the right-hand side, we get

$$\dot{V} = \frac{\lambda\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big((\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}})\boldsymbol{e}_L^\top - (\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}})\boldsymbol{e}_R^\top\Big)({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) \tag{46}$$

where the term $(\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}})\,\boldsymbol{e}_L^\top - (\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}})\,\boldsymbol{e}_R^\top$ can be rewritten as

$$\big(\boldsymbol{e}_R \times {}^0\underline{\boldsymbol{z}}\big)\,\boldsymbol{e}_L^\top - \big(\boldsymbol{e}_L \times {}^0\underline{\boldsymbol{z}}\big)\,\boldsymbol{e}_R^\top = [{}^0\underline{\boldsymbol{z}}]_\wedge \left(\boldsymbol{e}_R\boldsymbol{e}_L^\top - \boldsymbol{e}_L\boldsymbol{e}_R^\top\right) \tag{47}$$

Thanks to (17), $\boldsymbol{e}_R\boldsymbol{e}_L^\top - \boldsymbol{e}_L\boldsymbol{e}_R^\top$ can be written as

$$\boldsymbol{e}_R\boldsymbol{e}_L^\top - \boldsymbol{e}_L\boldsymbol{e}_R^\top = -[\boldsymbol{e}_L \times \boldsymbol{e}_R]_\wedge \tag{48}$$

Thereby, (46) becomes

$$\dot{V} = -\frac{\lambda\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top [{}^0\underline{\boldsymbol{z}}]_\wedge [\boldsymbol{e}_L \times \boldsymbol{e}_R]_\wedge ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) \tag{49}$$

Using again (17), one can show that

$$[{}^0\underline{\boldsymbol{z}}]_\wedge [\boldsymbol{e}_L \times \boldsymbol{e}_R]_\wedge ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) = -\,{}^0\underline{\boldsymbol{z}}^\top (\boldsymbol{e}_L \times \boldsymbol{e}_R)({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) \tag{50}$$

As a consequence, the time-derivative of the Lyapunov function candidate has for expression

$$\dot{V} = \frac{\lambda\epsilon\Big({}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)\Big)}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) \tag{51}$$

From (18), note that

$$\frac{\epsilon\Big({}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)\Big)}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} = \frac{-\text{sign}\Big({}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)\Big)\Big({}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)\Big)}{|\,{}^0\underline{\boldsymbol{z}}^\top(\boldsymbol{e}_L \times \boldsymbol{e}_R)|} = -1 \tag{52}$$

As a consequence, we finally get

$$\dot{V} = -\lambda({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) \tag{53}$$

It is strictly negative away from the desired configuration (asymptotic stability) and from the singularity case where the laser crosses the line passing by the two epipoles. Therefore, we have proven the asymptotic stability condition.

Now, it is easy to show that

$$({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) = 1 - ({}^0\underline{\boldsymbol{z}}^\top\,{}^0\underline{\boldsymbol{z}}^*)^2 \tag{54}$$

and that V develops as

$$V = 1 - {}^0\underline{\boldsymbol{z}}^\top\,{}^0\underline{\boldsymbol{z}}^* \tag{55}$$

from which we have

$$\dot{V} = -\lambda(1 + {}^0\underline{\boldsymbol{z}}^\top\,{}^0\underline{\boldsymbol{z}}^*)V \tag{56}$$

Unless ${}^0\underline{\boldsymbol{z}}$ and ${}^0\underline{\boldsymbol{z}}^*$ are diametrally opposed (which is an unrealistic case), the term $(1 + {}^0\underline{\boldsymbol{z}}^\top\,{}^0\underline{\boldsymbol{z}}^*)$ is strictly positive. As a consequence, V is exponentially decaying and, by this, we have proven the exponential stability. □

Proposition 2 (Stability condition of trajectory tracking) *The control in (33) is asymptotically stable if*

- *the laser does not cross the line passing through the two laser-image epipolar points*
- *and*

$$-\lambda(1 + {}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^*) + ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \omega^* < 0 \tag{57}$$

where $\omega^* = {}^0\underline{\boldsymbol{z}}^* \times {}^0\dot{\underline{\boldsymbol{z}}}^*$ *is the desired velocity of the laser beam direction.*

Proof: Let us consider the same Lyapunov function candidate as in Proposition 1 and its time derivative, now with a time-varying desired orientation of the laser (${}^0\dot{\underline{\boldsymbol{z}}}^* \neq \mathbf{0}$):

$$\dot{V} = -{}^0\underline{\boldsymbol{z}}^{*\top} {}^0\dot{\underline{\boldsymbol{z}}} - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \tag{58}$$

Inserting (27) and (33) into (58), now with a time-varying desired laser position, yields

$$\begin{aligned} \dot{V} \;=\; & \frac{\lambda\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big({}^0\underline{\boldsymbol{z}} \times \big(\boldsymbol{h}_L \times \boldsymbol{h}_R^* - \boldsymbol{h}_R \times \boldsymbol{h}_L^*\big)\Big) \\ & + \frac{\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big({}^0\underline{\boldsymbol{z}} \times \big(\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^* - \boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^*\big)\Big) \\ & - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \end{aligned} \tag{59}$$

We have already proven that the term on the first line, corresponding to the proportional feedback on a step response, yields exponential stability (Proposition 1). We now need to examine the two other terms, corresponding to the feedforward and to the proportional feedback on a time-varying response. Let us start with the one on the second line above and develop ${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^*)$ using (17)

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^*) = {}^0\underline{\boldsymbol{z}}^\top \dot{\boldsymbol{h}}_R^* \boldsymbol{h}_L - \underbrace{{}^0\underline{\boldsymbol{z}}^\top \boldsymbol{h}_L}_{0} \dot{\boldsymbol{h}}_R^* \tag{60}$$

With $\dot{\boldsymbol{h}}_R^* = \boldsymbol{e}_R \times {}^0\dot{\underline{\boldsymbol{z}}}^*$ and rearrangement of the mixed product, this becomes

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^*) = -\boldsymbol{h}_L \boldsymbol{h}_R^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \tag{61}$$

Doing the same for ${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^*)$ yields

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^* - \boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^*) = (-\boldsymbol{h}_L \boldsymbol{h}_R^\top + \boldsymbol{h}_R \boldsymbol{h}_L^\top) {}^0\dot{\underline{\boldsymbol{z}}}^* \tag{62}$$

which can be identified to

$$\begin{aligned} {}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^* - \boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^*) \;&=\; -(\boldsymbol{h}_R \times \boldsymbol{h}_L) \times {}^0\dot{\underline{\boldsymbol{z}}}^* \\ &=\; -\epsilon \|\boldsymbol{h}_R \times \boldsymbol{h}_L\| {}^0\underline{\boldsymbol{z}} \times {}^0\dot{\underline{\boldsymbol{z}}}^* \end{aligned} \tag{63}$$

Inserting the latter back into the last two lines of (59) develops and simplifies as

$$\begin{aligned} & \frac{\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big({}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^* - \boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^*\right)\Big) - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \\ = \; & \frac{\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \left(-\epsilon \|\boldsymbol{h}_R \times \boldsymbol{h}_L\| \, {}^0\underline{\boldsymbol{z}} \times {}^0\dot{\underline{\boldsymbol{z}}}^*\right) - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \end{aligned} \tag{64}$$

Simplifying by $\epsilon\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|$ yields now

$$\begin{aligned} & \frac{\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big({}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^* - \boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^*\right)\Big) - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \\ = \; & -\Big(({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top ({}^0\underline{\boldsymbol{z}} \times {}^0\dot{\underline{\boldsymbol{z}}}^*)\Big) - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \end{aligned} \tag{65}$$

Further development and rearrangement give

$$\begin{aligned} & \frac{\epsilon}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \Big({}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_L \times \dot{\boldsymbol{h}}_R^* - \boldsymbol{h}_R \times \dot{\boldsymbol{h}}_L^*\right)\Big) - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \\ = \; & -\left(-{}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^* {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^*\right) - {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \\ = \; & ({}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^* - 1) \, {}^0\underline{\boldsymbol{z}}^\top {}^0\dot{\underline{\boldsymbol{z}}}^* \end{aligned} \tag{66}$$

Using (27), it is possible to express ${}^0\dot{\underline{\boldsymbol{z}}}^*$ as

$${}^0\dot{\underline{\boldsymbol{z}}}^* = \omega^* \times {}^0\underline{\boldsymbol{z}}^* \tag{67}$$

The last two lines of (59) can be expressed finally as

$$\underbrace{-({}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^* - 1)}_{=V}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \omega^* \tag{68}$$

Thereby, the stability condition writes as

$$-\lambda\| {}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*\|^2 + ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \omega^* V < 0 \tag{69}$$

which can be factorized by V thanks to (54) and (55):

$$\left(-\lambda(1 + {}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^*) + ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \omega^*\right) V < 0 \tag{70}$$

and therefore we have proven that the stability condition writes as

$$-\lambda(1 + {}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^*) + ({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \omega^* < 0 \tag{71}$$

□

We can not prove a better result without further assumptions, because the second term of the left hand side of right hand side of (71) may be either positive or negative (Fig. 8) and thus contributes or opposes to the decrease of V. The first assumption we can make is that the tracking error remains small.

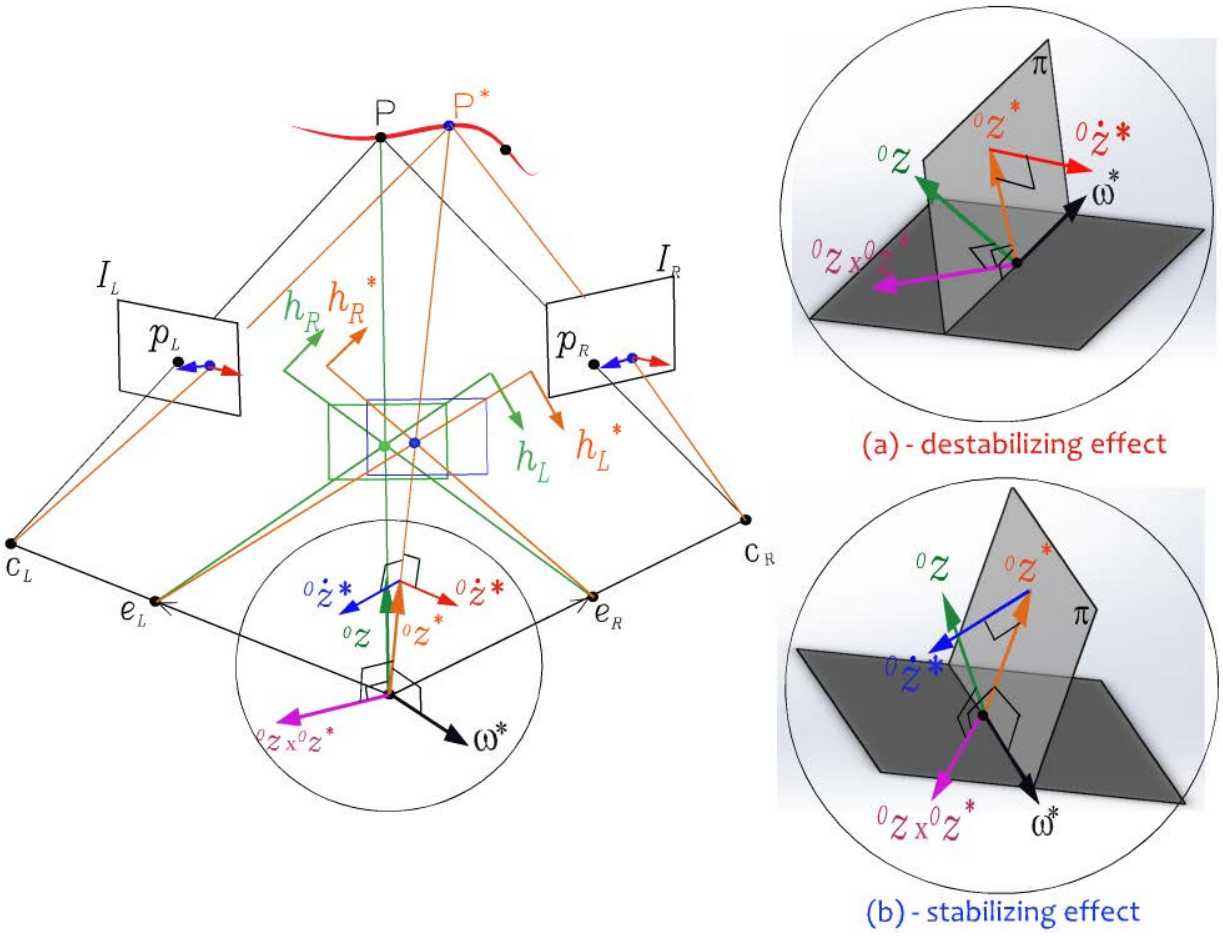


Figure 8: Representation of the stability and instability effects of the trajectory tracking.

Corollary 1 (Local asymptotic stability) *The control in (33) is locally asymptotically stable.*

Proof:

It appears clearly that, when the error ${}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*$ goes to 0, the second term of the left hand side of right hand side of (71) tends to -2λ, whereas the right-hand side term tends to 0 (provided finite $\boldsymbol{\omega}^*$). Therefore, the stability condition in Proposition 2 is satisfied at low errors and the control is locally asymptotically stable. □

In practice, the domain of asymptotic stability is practically speaking rather large because even when the angle θ between ${}^0\underline{\boldsymbol{z}}$ and ${}^0\underline{\boldsymbol{z}}^*$ is $\pi/10$, $1 + {}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^* (\approx 1.95)$ still dominates $\|({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)\| (\approx 0.31)$.

The second assumption is that the sampling frequency is high enough to allow for high proportional gain.

Corollary 2 (Gain tuning) *The control in (33) is globally asymptotically stable if the laser does not cross the line passing through the two laser-image epipoles and the proportional*

control gain λ satisfies

$$\forall {}^0\underline{\boldsymbol{z}},\ {}^0\underline{\boldsymbol{z}}^*, \omega^*, \lambda > \frac{\left|({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*)^\top \omega^*\right|}{(1 + {}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^*)} \tag{72}$$

In any case and without any assumption, mirroring the latter Corollary, keeping low the dynamics of the desired trajectory (ω^*) keeps the control stable.

Corollary 3 (Bounded dynamics) *The control in (33) is globally asymptotically stable if the laser does not cross the line passing through the two laser-image epipoles and the desired laser beam velocity is bounded by*

$$\forall {}^0\underline{\boldsymbol{z}},\ {}^0\underline{\boldsymbol{z}}^*, \lambda > 0, \|\omega^*\| < \lambda \frac{(1 + {}^0\underline{\boldsymbol{z}}^\top {}^0\underline{\boldsymbol{z}}^*)}{\|{}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*\|} \tag{73}$$

4.6 Discussion

4.6.1 Equivalence between 2D and 3D visual servoing

Reading carefully, the above proofs, the reader will be convinced that the control law we proposed is perfectly equivalent to

$$\boldsymbol{\omega} = \lambda({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{z}}^*) + ({}^0\underline{\boldsymbol{z}} \times {}^0\dot{\underline{\boldsymbol{z}}}^*) \tag{74}$$

The latter is the intuitive controller that would be derived in the 3D Euclidean space. Therefore, as in the case of visual servoing with respect to a sphere (Tatsambon Fomena and Chaumette, 2009), IBVS and PBVS are equivalent. The advantage of the proposed control is that it shows that is not necessary to perform calibration and triangulation in the Euclidean space, thanks to the trifocal constraint.

4.6.2 Singularities

However, the trifocal constraint we used is degenerate when the laser beam ${}^0\underline{\boldsymbol{z}}$ crosses the line between the two epipolar points in the virtual image (*i.e.* when the laser lies in the so-called trifocal plane made of the three centers of projection), because then $\boldsymbol{h}_R \times \boldsymbol{h}_L = 0$. In that case, neither the image-control nor the 3D one can be used. Indeed, the former becomes singular and the latter does not have any estimate for ${}^0\underline{\boldsymbol{z}}$ to be reconstructed.

The alternative to the 3D control would be to use the full trifocal tensor to reconstruct ${}^0\underline{\boldsymbol{z}}$, because it is not degenerate in the above-mentioned singular configuration. Following the proposed -methodology, one could extend the image-based control by differentiating with respect to time the trifocal tensor equations (5), thus yielding a constraint of the form

$$\begin{array}{lll} [\dot{\tilde{\boldsymbol{p}}}_R]_\times({}^0\underline{\boldsymbol{z}}_i\mathcal{T}_i^{jk})[\tilde{\boldsymbol{p}}_R]_\times & + & \\ [\tilde{\boldsymbol{p}}_R]_\times({}^0\dot{\underline{\boldsymbol{z}}}_i\mathcal{T}_i^{jk})[\tilde{\boldsymbol{p}}_R]_\times & + & \\ [\tilde{\boldsymbol{p}}_R]_\times({}^0\underline{\boldsymbol{z}}_i\mathcal{T}_i^{jk})[\dot{\tilde{\boldsymbol{p}}}_R]_\times & = & 0_{3\times3} \end{array} \tag{75}$$

However, this would require numerical matrix inversions, which we want to avoid in this paper.

Furthermore, the singular case is very unlikely to happen if the laser center of rotation does not lie in between the two real views. Notably enough, this constraint goes in the same way as the minimal footprint constraint for the system design, especially in the case of an endoscopic set-up.

4.6.3 Summary

The above results can be summarized by

- the proposed controller was shown to be equivalent in 2D and 3D, with a preference to the image-based formulation because one does not need to care for the sign of ${}^0\underline{z}$;
- the proposed controller does not require matrix calculus;
- the proposed controller is exponentially stable for a point-to-point task (except in a singular configuration rather easy to detect);
- it is only locally asymptotically stable for trajectory tracking. However, it turns globally asymptotically stable provided that high quality real-time hardware is available or that the desired trajectory has reasonable dynamics.

Note that the stability for trajectory tracking will not be experimentally tested in the sequel, because it turns out that path following is better suited to laser surgery than trajectory tracking, as we investigated it in the monocular case (Séon et al., 2014) (Séon et al., 2015).

5 Robustness Analysis in Simulation

5.1 Methodology

To analyze the robustness of the proposed control law with respect to calibration errors, we used the well-known formula of the fundamental matrix between the virtual view and the right camera

$$ {}^0\mathbf{F}_R = {}^R\mathbf{R}_\mathbf{0}^\top [\mathbf{R_0}]_\wedge \mathbf{K}_R^{-\mathbf{1}} \tag{76} $$

where $({}^R\mathbf{R_0}, \mathbf{R_0})$ is the pose of the right camera with respect to the laser pan-tilt manipulator and $\mathbf{K}_R$ contains the intrinsic parameters for the right camera. Trivially, the same applies to the left camera.

Thereby, a convenient way to apply calibration errors on a fundamental matrix is to apply disturbances on each component: extrinsic (relative pose) and intrinsic parameters.

The reference configuration is typical from an endoscopic set-up: the left (resp. right) camera is located 3.5 mm to the left (resp. right) of the laser manipulator, 10 mm above and

Error (pixel)	< 1	$\geq 1, < 5$	$\geq 5, < 20$	≥ 20	Singularity
Distribution	75%	$\approx 1\%$	$\approx 1\%$	23%	5%

Table 1: Distribution of the errors of convergence for 1000 simulations in presence of extrinsic calibration errors: 30 deg in rotation and 10 mm in translation.

5 mm forward. Their image planes are orthogonal to the laser beam in the zero-configuration. The two cameras are identical and their common intrinsic parameters are

$$\mathbf{K} = \begin{pmatrix} 400 & 0 & 320 \\ 0 & 400 & 240 \\ 0 & 0 & 1 \end{pmatrix} \tag{77}$$

The targeted tissue is located at 25 mm distance in front of the cameras.

Also, we tested the robustness of the control law to the geometric coherence of the desired configuration. Namely, do $\tilde{\boldsymbol{p}}_L$ and $\tilde{\boldsymbol{p}}_R$ correspond to an admissible 3D orientation of the laser? To do so, we simply added a random image offset on each feature.

Finally, in all simulation runs, we added a 2 pixel uniform noise to the tracking of each feature.

5.2 Results

Robustness to Intrinsic Parameters We added up to 30% relative error on the focal lengths and up to 100 pixels absolute error on the central point. Running the simulation 1000 times, never raised any unexpected behavior. This is not surprising since the reference is defined in the image, not in space.

Robustness to Extrinsic Parameters We rotated each camera around a random axis with a random angle up to 30 deg and translated them with up to 1 cm. The bi-modality exhibited by the distribution of the achieved errors shown in Table 1 suggests that extrinsic parameters have a reduced influence provided they respect some geometric constraints, to be further investigated. However, coming back to values more consistent with the assembly accuracy in an endoscope (5 deg and 1 mm), the distribution contains 100% of convergent runs.

Robustness to Geometric Incoherence The control is somehow less robust to the geometric incoherence. Actually, the control remains stable but converges in a local minimum, as depicted in Fig. 9, where the spot does neither reach the geometrically incoherent desired configuration (magenta 'x' cross) nor the original geometric configuration (red '+' cross). The system falls into a local minimum since the velocities of the micromanipulator converge to zero, up to noise. Note that this example is a dramatic one, since the difference between the coherent and incoherent configurations is larger than 20 pixels.

Error (pixel)	< 1	$\geq 1, < 5$	$\geq 5, < 20$	≥ 20	Singularity
Distribution	15%	62%	23%	0.1%	0.1%

Table 2: Distribution of the errors of convergence for 1000 simulations in presence of realistic noise and calibration errors.

Further geometric investigation is required to understand how the local minimum is "*chosen*" by the control law. Meanwhile, it is recommended to be very cautious in defining the desired reference through accurate matching between the left and right spots. For instance, 15% of our 1000 simulation runs with a geometric incoherence below 4 pixels converged within less than 1 pixel, and an additional 62% to less than 5 pixels.

Robustness to all Disturbances Simultaneously We mixed all disturbances with up to 5 deg rotation and 1 mm translation of each cameras, 20% relative error on the focal lengths, 30 pixel deviation of the central point, 4 pixel deviation in each image from a geometrically coherent desired configuration and 1 pixel detection error. The statistics of 1000 simulation runs are given in Table 2.

6 Experimental Validation

6.1 Experimental Set-up

The proposed approach was validated on the experimental set-up shown in Fig. 10[2]. It consists of two CCD cameras (frame rate of 25 images per second for a resolution of 640×480 pixels), a laser source (laser pointer), a fixed mirror, an actuated mirror (the S-334 from Physical Instruments Inc[3]). The latter is based on a parallel-kinematics design with three coplanar axes and a fixed pivot point. Thereby, these axes give two rotations (α and β) which are characterized by a large bandwidth (1kHz), a resolution of 0.2 μrad, a linearity of 0.05 %, and a motion range of $[-0.026\ rad, +0.026\ rad]$.

The current functioning mode of this experimental set-up is as follows: the operator defines a desired position $\tilde{\boldsymbol{p}}^*_{L}$ (using a mouse-click) to reach a point in the left image, and by using the fundamental matrix ${}^{L}\mathbf{F}_{R}$ between the left camera and the right camera, it is possible to find the corresponding desired position $\tilde{\boldsymbol{p}}^*_{R}$ in the right image by searching through the epipolar line ($\tilde{\boldsymbol{p}}_{R} = {}^{L}\mathbf{F}_{R}\tilde{\boldsymbol{p}}_{L}$). Note that this requires a rather good estimation of ${}^{L}\mathbf{F}_{R}$, but this is the standard F-matrix estimation problem, available in any good library (*e.g.* OpenCV).

The time-varying positions (i.e., current positions) of the laser spot, in both images, are tracked simultaneously using ViSP (Visual Servoing Platform) (Marchand et al., 2005). The developed controller was tested in two different scenarios to assess its behaviors (accuracy, convergence, decoupling, etc.). The first case of study consists of the control of the laser

[2]A video presenting this experimental set-up can be seen in the Extension 1 **(refer to Appendix (8.1))**.
[3]`http://www.physical-instruments.fr`

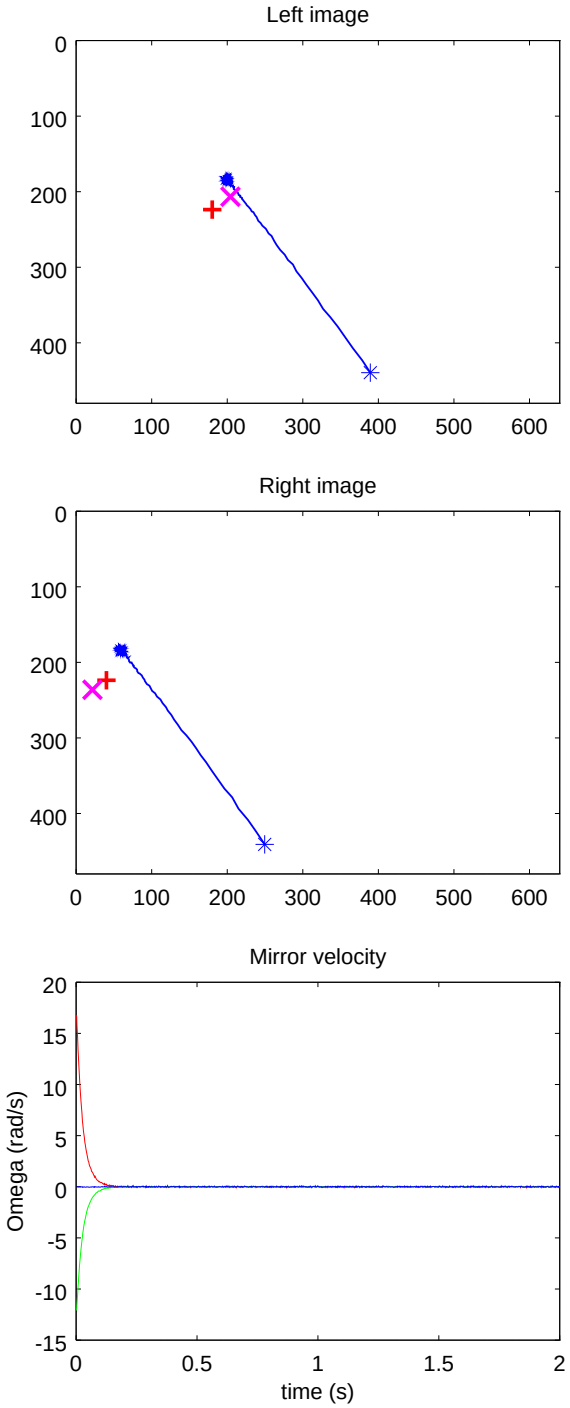


Figure 9: Robustness to geometric incoherence. Example of a discrepancy between a geometrically coherent desired location of the spot in each image (red '+') and a geometrically incoherent one (magenta 'x') obtained by some 20 pixel offset. From top to bottom: trajectory in the left and right image, velocities of the mirror with respect to time.

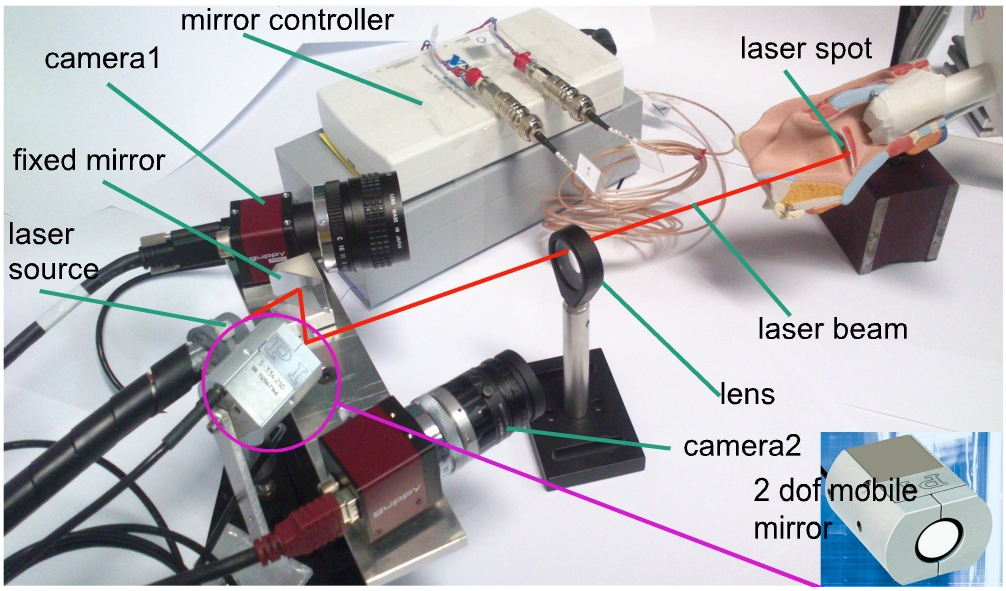


Figure 10: Photography of the experimental set-up.

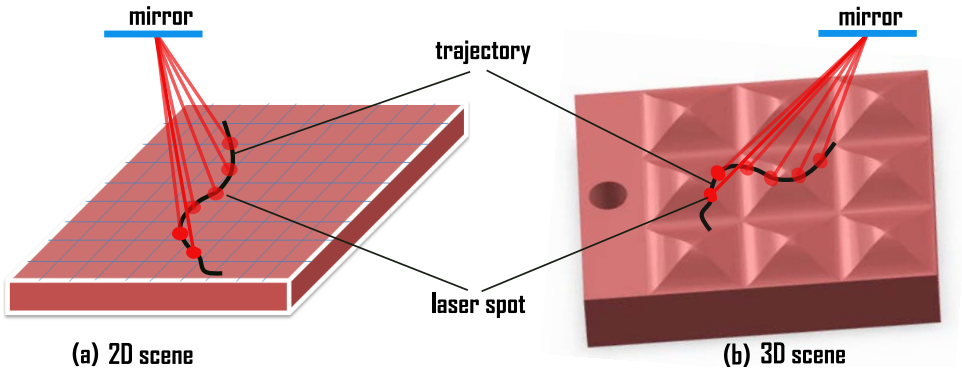


Figure 11: Validation scenarios: (a) and (b) represent the 2D and 3D surfaces used as targeted surfaces, respectively.

spot displacement over a simple planar scene (Fig. 11(a)) while the second uses a home-made 3D shape surface with 5mm high hills on which the laser spot evolves during the control (Fig. 11(b)). Obviously, the latter is much closer to the intended use of this work i.e., vocal folds laser surgery. We are aware that it lacks the consideration of the variation of vocal cords during a surgical intervention.

6.2 Calibration Issues

It is important to underline that the experimental validations described in this section are obtained without any accurate camera calibration (intrinsic matrix $\mathbf{K}_{3\times3}$) nor accurate Euclidean (*hand-eye*) calibration (i.e., left camera/mirror and right camera/mirror poses). In-

stead, we just need a weak calibration: image-to-laser fundamental matrices. Furthermore, these matrices do not even need to be accurate

- ${}^{0}\mathbf{F}_{L} = [{}^{0}\mathbf{t}_{L}]_{\wedge}{}^{0}\mathbf{R}_{L}\mathbf{K}^{-1}$, the fundamental matrix between the virtual view associated to the laser beam and the left view, was approximated (manually) by a translation vector ${}^{0}\mathbf{t}_{L} = (1, -0.1, 38)^{\top}$ *mm* and a rotation matrix ${}^{0}\mathbf{R}_{L} = \mathbf{I}_{3\times 3}$. Similarly, ${}^{0}\mathbf{F}_{R} = [{}^{0}\mathbf{t}_{R}]_{\wedge}{}^{0}\mathbf{R}_{R}\mathbf{K}^{-1}$ was approximated by a translation vector ${}^{0}\mathbf{t}_{L} = (1, -20, 38)^{\top}$ *mm* and a rotation matrix ${}^{0}\mathbf{R}_{R} = \mathbf{I}_{3\times 3}$;

  This is obviously rather far away from the configuration shown in Fig. 10.

- $\mathbf{K}_{3\times 3}$ contains a coarse estimation of both camera's intrinsic parameters

$$\mathbf{K} = \begin{pmatrix} 1500 & 0 & 320 \\ 0 & 1500 & 240 \\ 0 & 0 & 1 \end{pmatrix} \tag{78}$$

6.3 Validation on a 2D Surface

Fig. 12 shows an image sequence captured by the left and right cameras during an automatic control of the laser spot displacement in a 2D surface[4]. More precisely, Fig. 12(l_1) illustrates, simultaneously in $\mathbf{I}_L$ and $\mathbf{I}_R$, the achievement of the first subtask i.e., the steering of laser between an initial position $\tilde{\boldsymbol{p}}_0$ and a first desired position $\tilde{\boldsymbol{p}}_{\mathbf{1}}{}^{*}$. When, the laser reaches the desired position, the user defines by a mouse-click the next desired position which the laser must attain. Fig. 12(l_1)-(l_4) represent the others subtasks performed by the controller. It can be seen that the laser spot reached the desired position with good accuracy for all turns. In the zoom over Fig. 12 shown in Fig. 13, the laser spot trajectory, in both images, is a straight-line for a point-to-point displacement (step response) without any path planning phase. So, we can observe that the difference between the ideal trajectory (straight-line) and the trajectory tracked by the laser spot is virtually superimposed. The small difference between the two trajectories can be explained by a problem of tracking robustness (the laser spot size, thus the coordinates of its gravity center, moves slightly during the control) and the limited effects of the rather strong calibration errors. This result means that our controller takes the shortest path both in images and space which is the best achievable performance without having any fine estimation of the different calibration matrices. Fig. 14(a) and 14(b) show the regulation to zero of the error $\mathbf{e} = \begin{pmatrix} \tilde{\boldsymbol{p}}_L - \tilde{\boldsymbol{p}}_L^{*} \\ \tilde{\boldsymbol{p}}_R - \tilde{\boldsymbol{p}}_R^{*} \end{pmatrix}$ in the left and right images *versus* the number of iterations i, respectively. From these figures, it can also be verified that the image errors follow exponential decay to reach the successive desired positions and that the control is perfectly decoupled (same exponential rate on all components, straight lines in the image). The overall task is accomplished with great accuracy in a dozen iterations where each one takes approximately 0.06 second (the camera frame rate 25 fps) and a gain of $\lambda = 3$.

[4]A video showing this validation test can be seen in the Extension 2 **(refer to Appendix (8.1) for more details)**

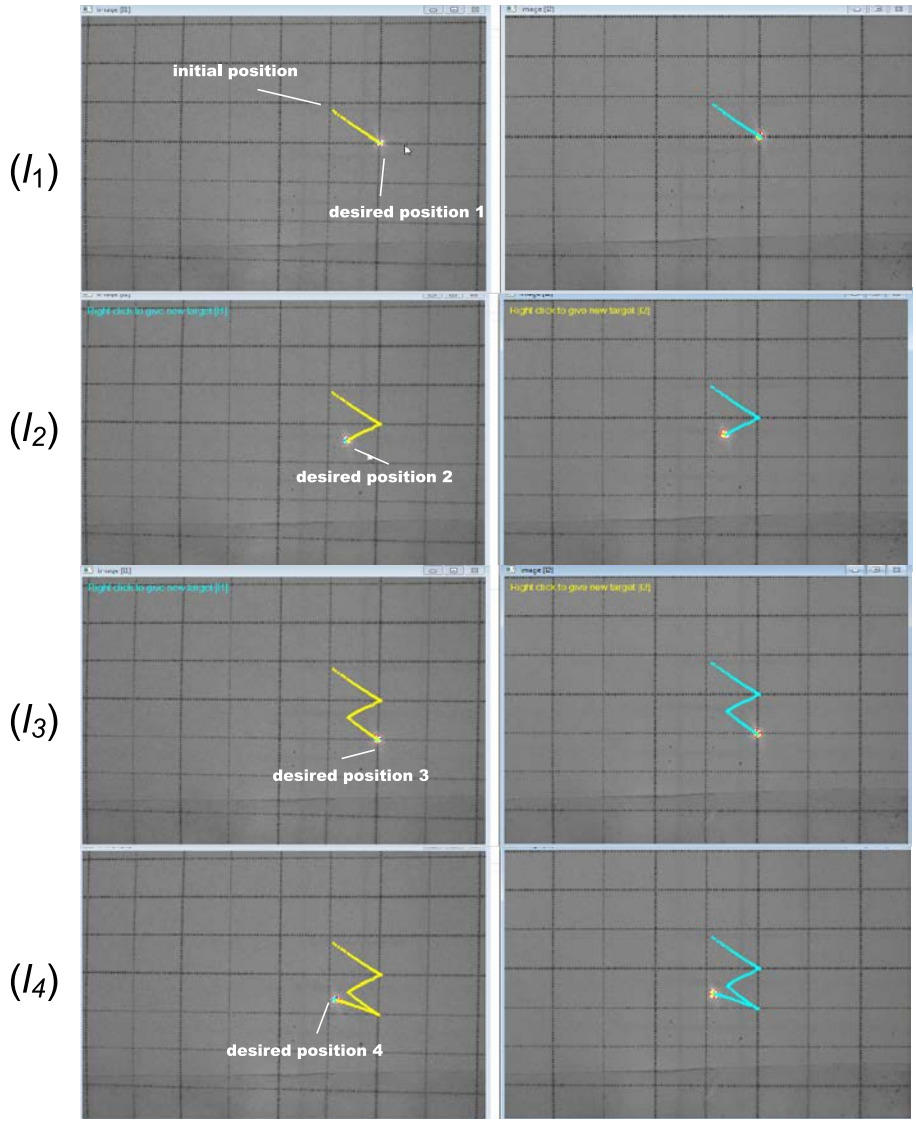


Figure 12: Image sequence taken from both cameras (the first column shows the left camera and the second column shows the right camera) which illustrates the laser spot displacement during the control. It moves from an initial position $\tilde{\boldsymbol{p}}_{\mathbf{0}}$ to successive desired positions (from $\tilde{\boldsymbol{p}}_{\mathbf{1}}{}^{*}$ to $\tilde{\boldsymbol{p}}_{\mathbf{4}}{}^{*}$) without any path planning phase.

The latter is determined empirically in order to find the best comprise between accuracy and rapidity. Concerning the accuracy of the experimental validation in a 2D target, the RMS (Root Mean Square) of the steady-state error is 0.38 pixels (i.e., 47 μm) with a standard deviation (STD) of 1.8 pixels (i.e., 230 μm). According to the vocal folds laser surgery specifications, the admissible error on the laser control is 200 μm. It can be highlighted that the proposed controller is then in line to the vocal folds laser surgery recommendations.

Concerning the decoupling aspect of the controller, according to the Fig. 14 which il-

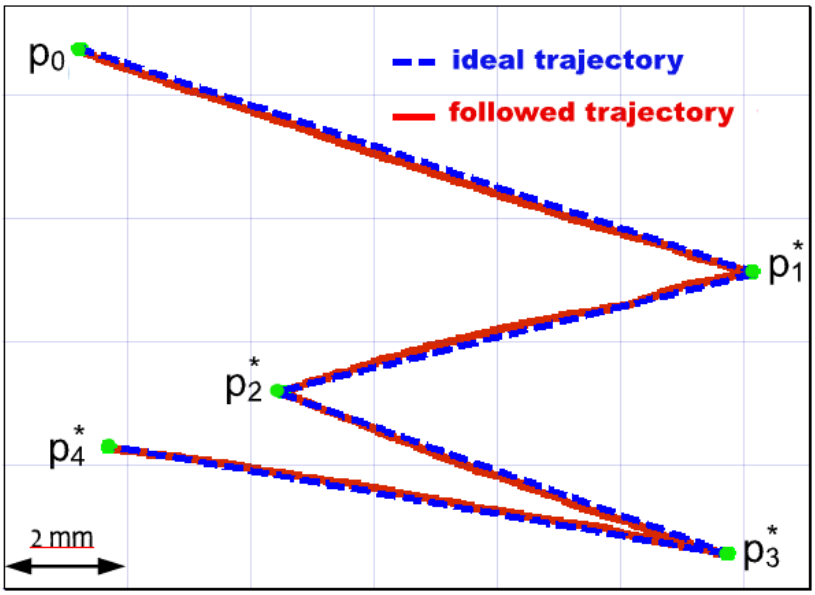
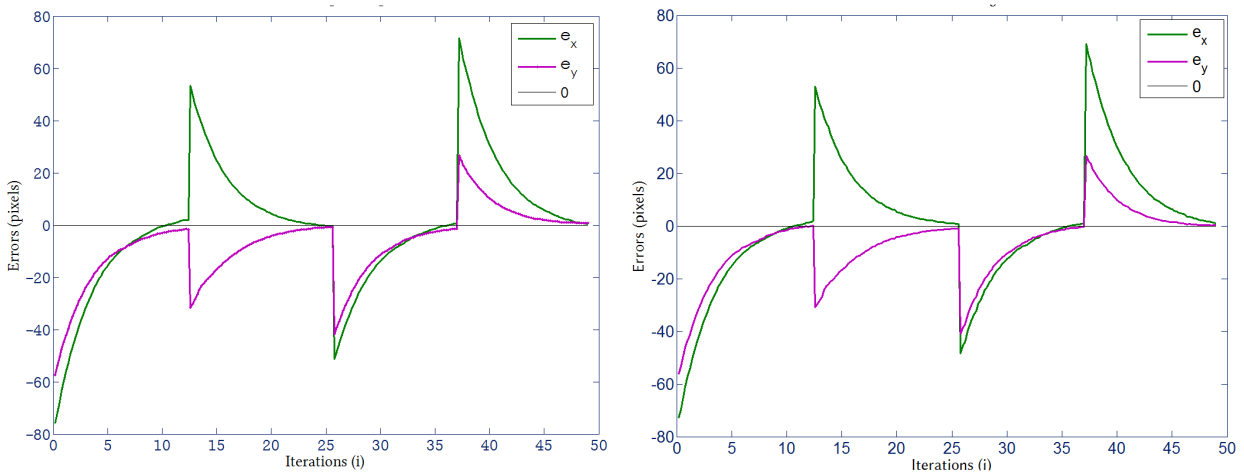


Figure 13: Superposition of the performed trajectory by the laser spot in $\mathbf{I}_L$ and an ideal trajectory

(i.e., a point-to-point).



(a) Left camera.

(b) Right camera.

Figure 14: Image errors (in pixels) *vs.* number of iterations i: (a) left camera and (b) right camera.

lustrates a convergence of the all the components of the error in the same rate (here, in an exponential decrease). This behavior can be seen also on the trajectories in the control space which are in straight-line Fig. 13. For completeness sake, Fig. 15 shows the joint velocities (i.e., $\dot{\mathbf{q}}_1$ and $\dot{\mathbf{q}}_2$) of the steering mirror during the control but this only shows the good quality of the industrial manipulator we used. Table 3 gathers the various time consumptions

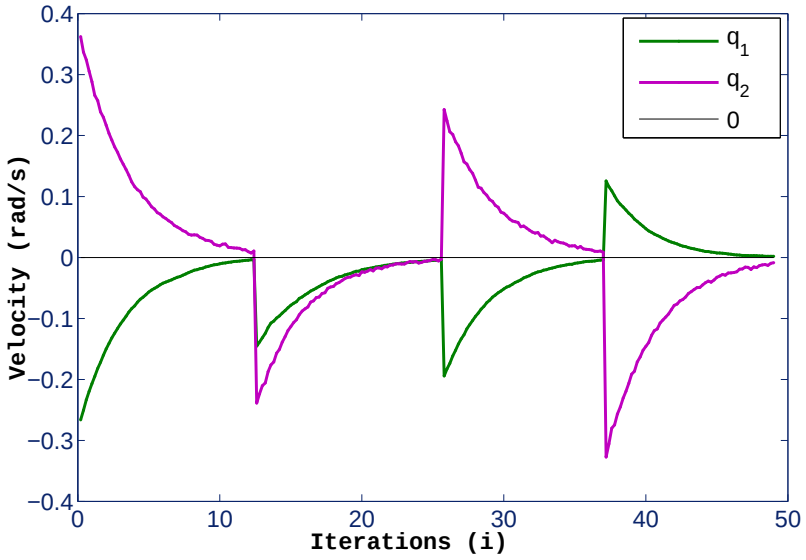


Figure 15: Joint velocities $\dot{\mathbf{q}}_1$ and $\dot{\mathbf{q}}_2$ *vs.* number of iterations i.

Table 3: Time-consuming of the main tasks of the process.

task	time (ms)
grabbing image	40
laser spot tracking	0.65
controller	0.001
sending control (USB 2.0)	0.989
others (displays, variables declaration, etc.)	0.176

recorded during the experiments. It shows that the proposed controller takes only 0.001 ms, far less than the image grabbing time and the communication with the manipulator. We thus presume that using efficient hardware (high-speed camera and control electronics) would allow for reaching control frequencies above 1kHz.

6.4 Validation on a 3D Surface

The laser steering is also validated on a scene Fig. 11(b) that is closer to the final application in order to study the performance of the controller in realistic conditions of use. Thereafter, the validation scenario remains the same as in the case of 2D surface: the user clicks successively on the image to define the different desired positions. Obviously, the idea is to observe the behavior of the laser spot displacement in this object for different directions and with trajectories going from one vertex to another. It is also important to underline that this experiment was performed with the same rough calibration as the previous one.

The results of this second scenario are shown in Fig. 16[5]. The first and second columns

[5]This experimental validation can be seen in the Extension 3 (refer to Appendix (8.1) for more details)

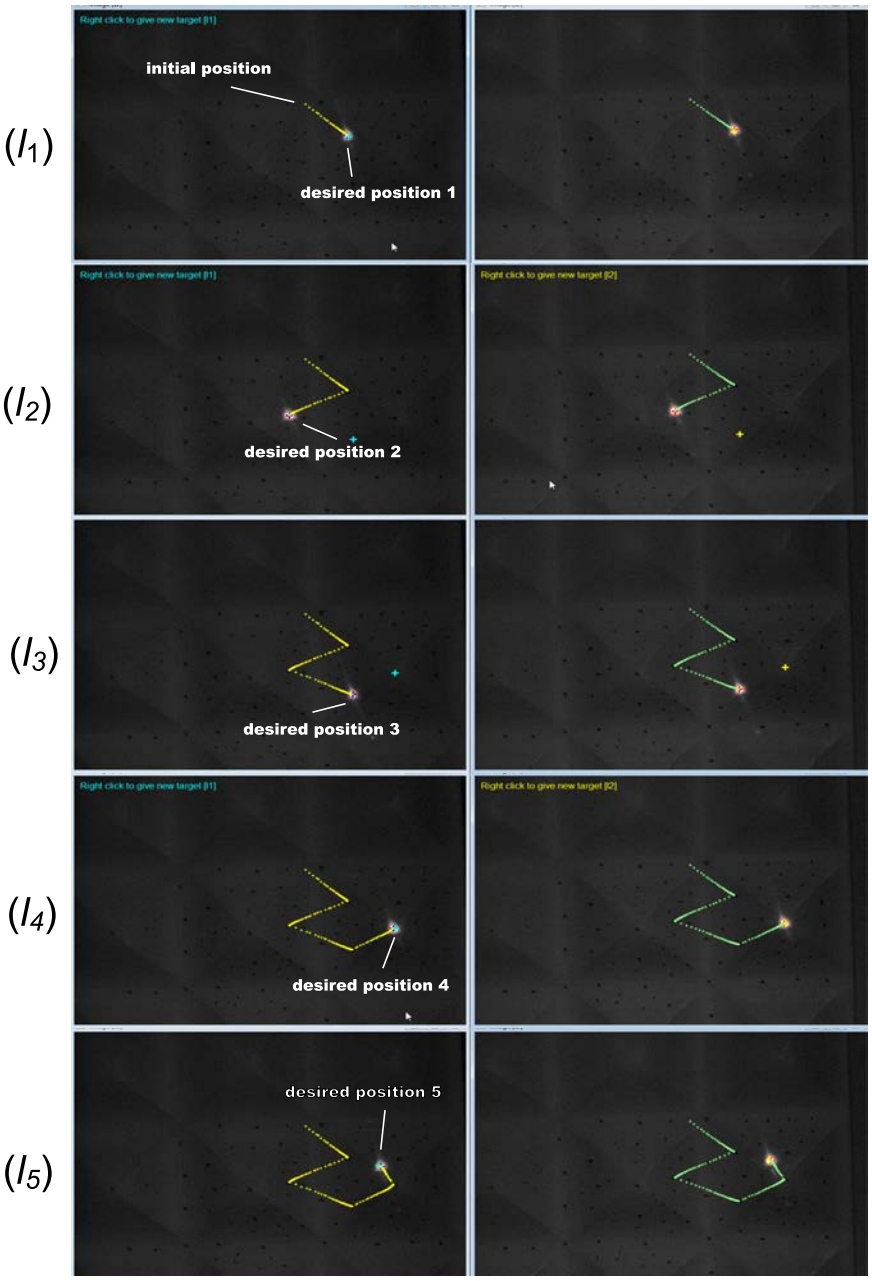


Figure 16: Images sequence captured during the experiments in a 3D scene. As in the first experiment, the laser spot moves from an initial position $\tilde{\boldsymbol{p}}_{\mathbf{0}}$ to successive desired position $\tilde{\boldsymbol{p}}_{\mathbf{1}}{}^{*}$ to $\tilde{\boldsymbol{p}}_{\mathbf{5}}{}^{*}$ without any path planning phase.

represent the images captured by the left camera and the right camera, respectively. Fig. 16(l_1) shows the achievement of the first subtask where the laser spot moves from the initial position $\tilde{p}$ to the different desired positions $\tilde{\boldsymbol{p}}_{\mathbf{1},\ldots,\mathbf{5}}{}^{*}$. The achieved trajectory is shown in Fig. 17. It can be highlighted that the laser spot trajectory is again very close to the straight-line

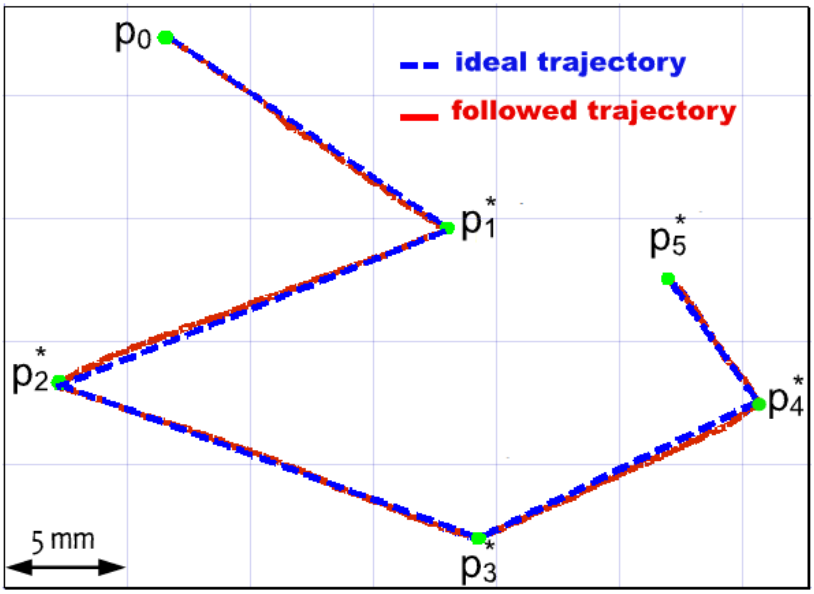


Figure 17: Superposition of the laser spot trajectory in $\mathbf{I}_L$ with an ideal point-to-point trajectory.

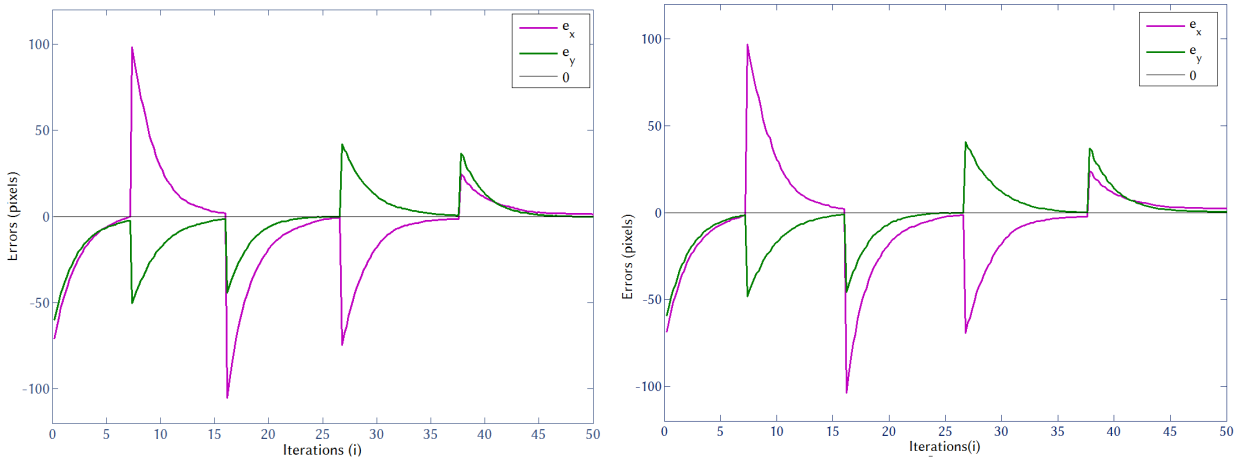
trajectory. This is performed without any prior knowledge nor 3D estimation (e.g., 3D reconstruction or depth computation) of the used surface. However, there is a small gap between the two trajectories. This is due to the fact that the shape and the size of the laser spot changes depending on the surface. For instance, it grows larger (it becomes elliptical) when it moves on a not parallel plane to the image plane. In the same case, the errors regulation to zero in the left image and right image are shown in Fig. 18(a) and (b), respectively. Similarly to 2D scene experiments, it can be observed that the controller presents good exponential convergence and decoupling behaviors.

As regards the accuracy of the developed controller in a 3D target, the following values are computed: the RMS of the steady-state error is 1.66 pixels (i.e., 210μm) with a STD of 2.07 pixels (i.e., 259 μm).

6.5 Preliminary validation on human cadaver

A first stage of the validation tests of the controller, in clinical conditions, is performed on a human cadaver with the presence of the surgeon, using the prototype developed jointly within the μRALP project. Instead of the PI mirror (Fig. 10), a preliminary prototype of a home-made 2 dof small mirror of $8 \times 8 \times 12$ mm^3 (Fig. 19) (Rabenorosoa et al., 2014) was used. The mirror is integrated on the endoscope tip (of 18 mm of diameter) which embeds two color miniature CCD cameras[6], and a visible red fiber laser. The various fundamental matrices were estimated roughly before the tests. The protocol of use is the following, when the laser spot tracker is initialized, the surgeon defines the desired positions in both images $\mathbf{I}_L$

[6]http://www.misumi.com.tw



(a) Left camera. (b) Right camera.

Figure 18: Image errors (in pixels) *vs.* number of iterations i (3D surface): (a) left camera and (b) right camera.

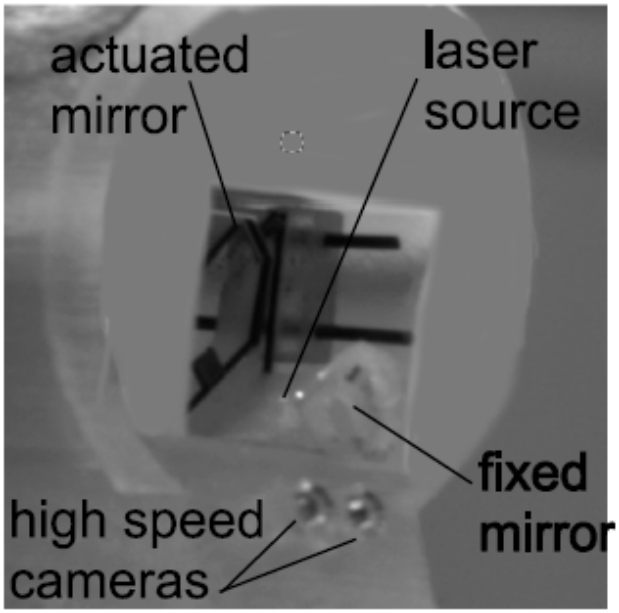


Figure 19: Photography of the first version of the developed endoscope tip.

and $\mathbf{I}_R$ using a pen tablet through the Surgeon-Robot Interface proposed in (Dagnino et al., 2012). During the visual servoing process, the surgeon can update continuously the desired positions to manually follow a path. Figure 20 shows some images of the larynx captured during the validation test. This realistic experimental validation (during cadaver trials) have demonstrated that the controller presents satisfactory characteristics with respect to the vocal folds surgery features. The accuracy remains interesting: RMS of the steady-error

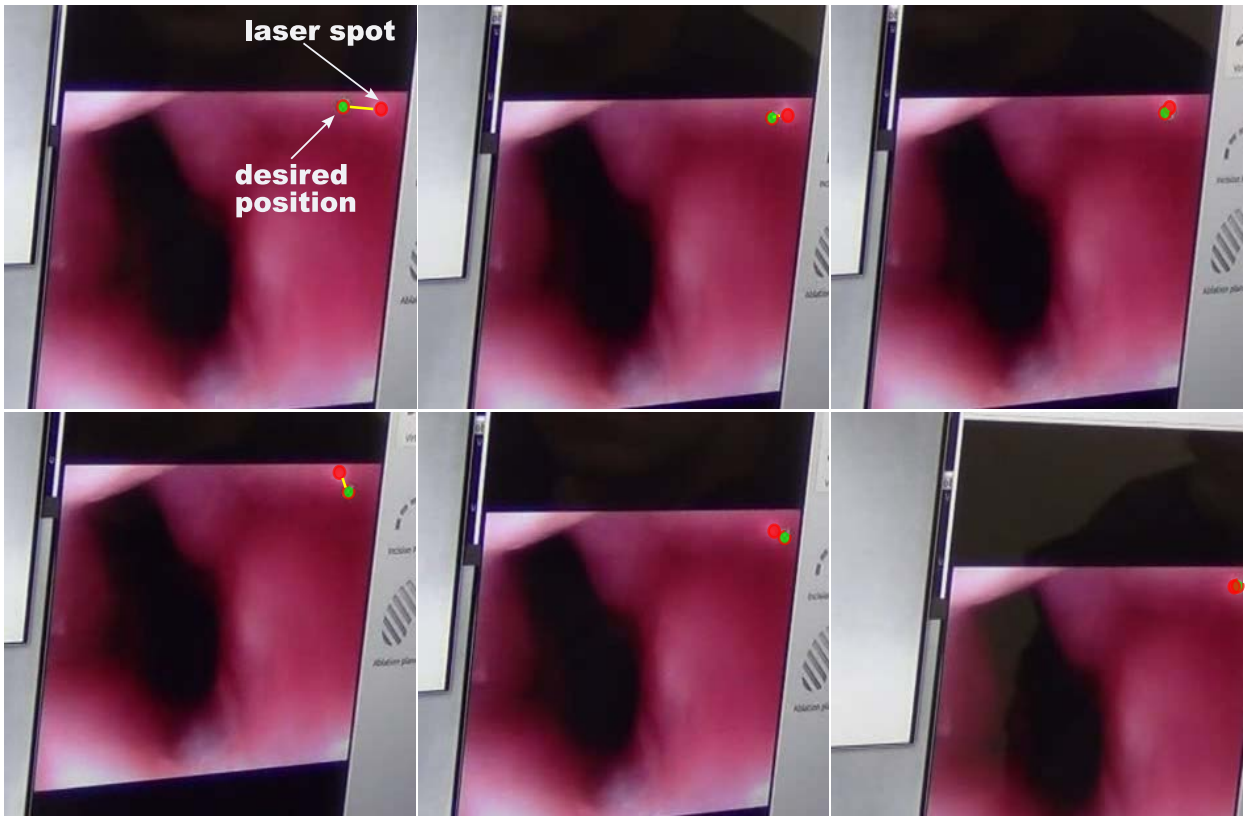


Figure 20: Validation of the controller during cadaver trials.

of 1.20 pixels (210 μm) with and STD of 1.54 pixels (269.5 μm) even using a home-made prototype as on a desktop and using a high quality commercial mirror. Concerning the convergence rate of the controller during the cadaver trial, the trifocal-based control law converges in fifteen iterations using the same gain $\lambda = 3$ (the same used for the testbench experiments). The stability of the controller has also been shown during these validations. The controller was held in check only when the laser spot tracking algorithm has failed especially because of the presence of high specular in the vocal folds tissues.

7 Conclusion and Future Work

In this paper, a visual servoing control law based on the use of an adapted formulation of the trifocal constraint has been presented in order to achieve an automatic control of the laser spot displacements. Laser surgery (e.g., laser microphonosurgery) often presents this scenario, where a camera observes the laser spot displacement (scanning) on a 3D surface, but we identified several other potential application cases.

More precisely, it was shown that using trifocal geometry simplifies the eye-to-hand visual control law of a pan-tilt laser. Namely, the control could be derived at hand, without any interaction matrix inversion nor any need for strong Euclidean calibration of the system.

Moreover, explicit stability conditions could be established (which is not so often in visual servoing), both for point-to-point tasks or for trajectory tracking.

The experimental results have demonstrated that the proposed control law was accurate, fully decoupled (both image trajectories were straight) with an exponential decay of the image errors and also robust to rather large errors on the laser-image fundamental matrices.

The next stages of this work will involve adapting the described hardware for real laser surgery applications. This will concern essentially the use of a high speed ($\geq$ 200 frames/second) camera which will be connected to two fiber bundles through miniatures GRIN lenses as well as the finalization of the home-made prototype. Further theoretical developments are also expected in order to replace trajectory tracking by path following in the proposed trifocal set-up.

Acknowledgment

The authors wish to thank Dr. Peter Sturm from the INRIA Grenoble Rhône-Alpes (STEEP research team), France, as well as the reviewers for all their comments allowing to improve this paper.

This work was supported by μRALP, the EC FP7 ICT Collaborative Project no. 288663 (http://www.microralp.eu), and by ACTION, the French ANR LABEX no. ANR-11-LABX-01-01 (http://www.labex-action.fr).

8 Appendix

8.1 Index to Multimedia Extensions

Extension	Type	Description
1	video	video showing the global view of the experimental set-up used in these work
2	video	video showing an example where the controller performs a successive simple tasks in both 2D scene.
3	video	video showing an example where the controller performs a successive simple tasks in unknown 3D scene.

8.2 Notation

Notations	Description
"." and $\times$	dot product and vectorial product, respectively
$[\ .\ ]_\times$	skew-symmetric cross product matrix
$\mathbf{I}_L$ and $\mathbf{I}_R$	left image and right image, respectively
$\tilde{\boldsymbol{p}}_L$ and $\tilde{\boldsymbol{p}}_R$	2D homogeneous points projected in $\mathbf{I}_L$ and $\mathbf{I}_R$, respectively
${}^0\mathbf{P}$	3D homogeneous point in the world frame ${}^0R$
$\tilde{\boldsymbol{p}}$	projection of ${}^0\mathbf{P}$ in the image $\mathbf{I}$
$\tilde{\boldsymbol{p}}^*$	desired image position
$\dot{\tilde{\boldsymbol{p}}}^*$	feed-forward term in case of trajectory tracking
$\boldsymbol{c}_n$	optical center of camera n
Φ_n	image plane of camera n
$(\boldsymbol{c}_n \boldsymbol{c}_m)$	baseline between camera n and camera m
$\boldsymbol{e}_n$	epipolar point of camera n
$(\boldsymbol{e}_n \boldsymbol{p}_n)$	epipolar line in the image n
${}^i\mathbf{F}_j$	fundamental matrix between the cameras i and j
$\mathcal{T}$	a $3 \times 3 \times 3$ trifocal tensor
$\mathbf{T}_i$	one of the three 3×3 matrix of $\mathcal{T}$
${}^0\underline{\boldsymbol{z}}$	z direction of the laser beam from the mobile mirror towards the vocal fold
${}^0\underline{\boldsymbol{z}}^*$	desired direction of the laser beam
${}^0\dot{\underline{\boldsymbol{z}}}$	time-derivative of ${}^0\underline{\boldsymbol{z}}$
d	distance between the mobile mirror and the vocal fold
Z	depth information of the scene
$\boldsymbol{h}_R$ and $\boldsymbol{h}_L$	normal vectors to the planes π_1 and π_2, respectively
$\boldsymbol{h}_L^*$ and $\boldsymbol{h}_R^*$	reference vectors of $\boldsymbol{h}_L$ and $\boldsymbol{h}_R$, respectively
$\dot{\boldsymbol{h}}_L^*$ and $\dot{\boldsymbol{h}}_R^*$	time-derivative of $\boldsymbol{h}_L$ and $\boldsymbol{h}_R$, respectively
$\dot{\tilde{\boldsymbol{p}}}_L$ and $\dot{\tilde{\boldsymbol{p}}}_R$	time-derivative of $\tilde{\boldsymbol{p}}_L$ and $\tilde{\boldsymbol{p}}_R$, respectively
$\boldsymbol{\omega}$	rotation velocity of the mirror
λ and κ	scalars
V	Lyapunov candidate function
$\dot{V}$	time-derivative of the Lyapunov candidate function
α and β	rotation dofs of the mobile mirror

References

F. Alkhalil and C. Doignon. Stereo visual servoing with decoupling control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1671–1676, 2012.

A. Aloy and M. Grasl. Endoscopy of larynx and trachea with rigid laryngo-tracheoscopes under superimposed high-frequency jet ventilation (shfjv). *Endoscopy*, Ed. Somchai Amornyotin, ISBN 978-953-51-1071-2, April, 2013. doi: 10.5772/52996.

N. Andreff and P. Martinet. Vision-based self-calibration and control of parallel kinematic mechanisms without proprioceptive sensing. *Intelligent Service Robotics*, 2(2), pages 71–80, 2009.

N. Andreff, B. Espiau, and R. Horaud. Visual servoing from lines. *International Journal of Robotics Research*, 21(8), pages 679–700, 2002.

N. Andreff, T. Dallej, and P. Martinet. Image-based visual servoing of a gough-stewart parallel manipulator using leg observations. *International Journal of Robotics Research*, 26(7), pages 677–687, 2007.

N. Andreff, S. Dembélé, B. Tamadazte, and Z. E. Hussnain. Epipolar geometry for vision-guided laser surgery. In *International Conference on Informatics in Control, Automation and Robotics*, pages 1–6, Iceland, 2013.

F. Arai, K. Yoshikawa, T. Sakami, and T. Fukuda. Synchronized laser micromanipulation of multiple targets along each trajectory by single laser. *Applied Physics Letter*, 85, pages 4301–4303, 2004.

P. Avanzini, E. Royer, B. Thuilot, and P. Martinet. A global decentralized control strategy for urban vehicle platooning using monocular vision and a laser rangefinder. In *International Conference on Control, Automation, Robotics and Vision*, pages 291–296, 2008.

H.M. Becerra and C. Sagüés. Exploiting the Trifocal Tensor in Dynamic Pose Estimation for Visual Control. *IEEE Transactions on Control Systems Technology*, 21 (5), pages 1931-1939, 2013.

H.M. Becerra and C. Sagüés. A robust control scheme based on the trifocal tensor. In *Springer Tracts in Advanced Robotics*, pages 45–68. Springer International Publishing, 2014.

H.M. Becerra, G. López-Nicolás, and C. Sagüés. A sliding-mode-control law for mobile robots based on epipolar visual servoing from three views. *IEEE Transactions on Robotics*, 27 (1), pages 175–183, 2011.

S. Benhimane and E. Malis. Homography-based 2d visual tracking and servoing. *The International Journal of Robotics Research*, 26(7), pages 661–676, 2007.

J. Courbon, Y. Mezouar, and P. Martinet. Indoor navigation of a non-holonomic mobile robot using a visual memory. *Autonomous Robots*, 25(3), pages 253-266, 2008.

F. Chaumette and S. Hutchinson. Visual servo control, part 1 : Basic approaches. *IEEE Robotics and Automation Magazine*, 13(1), pages 82–90, 2006.

F. Chaumette and S. Hutchinson. Visual servo control, part 2 : Advanced approaches. *IEEE Robotics and Automation Magazine*, 14(1), pages 109–118, 2007.

L. Coutard, F. Chaumette, and J. M. Pflimlin. Automatic landing on aircraft carrier by visual servoing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2843–2848, 2011.

H. Cui, Z. Xiao, J. Dong, Y. Chen, and Z. Hou, G. Zhao. Towards fast femtosecond laser micromachining of glass, effect of deposited energy. In *International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pages 2091–2094, 2013.

G. Dagnino, L. S. Mattos, and D. G. Caldwell. New software tools for enhanced precision in robot-assisted laser phonomicrosurgery. In *34th Annual International Conference of the Engineering in Medicine and Biology Society*, doi: 10.1109/EMBC.2012.6346547, 2012.

D. Dedieu, V. Cadenat, and P. Soueres. Mixed camera-laser based control for mobile robot navigation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1081–1086, 2000.

D.F. Dementhon and L.S. Davis. Model-based Object Pose in 25 Lines of Code. International Journal of Computer Vision, 15(1-2), pages 123–141, 1995.

H.E. Eckel, Sh. Berendes, M. Damm, and J.P. Klusmann. Suspension laryngoscopy for endotracheal stenting. *Laryngoscope*, 113, pages 11–15, 2003.

B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3), pages 313–326, 1992.

M. Fridenfalk and G. Bolmsjö. Design and validation of a universal 6d seam-tracking system in robotic welding using arc sensing. *Advanced Robotics*, 18, pages 1–21, 2004.

R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge, United Kingdom, 2nd edition, ISBN: 0521540518 2006.

J. Hespanha, Z. Dodds, G. D. Hager, and A. S. Morse. What can be done with an uncalibrated stereo system? In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1366–1372, 1998.

Y. Huang, Y. Xiao, P. Wang, and M. Li. A seam-tracking laser welding platform with 3d and 2d visual information fusion vision sensor system. *International Journal of Advanced Manufacturing Technology*, 67, pages 415–426, 2012.

S. Huh, D.H. Shim, and H. Kim. Integrated navigation system using camera and gimbaled laser scanner for indoor and outdoor autonomous flight of UAVs. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3158–3163, 2013.

S. Hutchinson, G.D. Hager, and P.I. Corke. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5), pages 651–670, 1996.

G.J. Jako. Laryngoscope for microscopic observation, surgery and photography. *Arch Otolaryngol*, 91, pages 196–199, 1970.

D. Khadraoui, C. Debain, R. Rouveure, P. Martinet, P. Bonton, and J. Gallice. Vision-based control in driving assistance of agricultural vehicles. *International Journal of Robotics Research*, 17(10), pages 1040–1054, 1998.

A. Krupa, C. Doignon, J. Gangloff, and M. de Mathelin. Combined image-based and depth visual servoing applied to robotized laparoscopic surgery. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 323–329, 2002.

A. Krupa, G. Fichtinger, and G.D. Hager. Real-time motion stabilization with b-mode ultrasound using image speckle information and visual servoing. *International Journal of Robotics Research, IJRR*, 28(10), pages 1334–1354, 2009.

B. Lamiroy, B. Espiau, N. Andreff, and R. Horaud. Controlling robots with two cameras: How to do it properly. In *IEEE International Conference on Robotics and Automation*, San Francisco, USA, pages 200-206, 2000.

V. Lippiello, B. Siciliano, and L. Villani. Position-based visual servoing in industrial multirobot cells using a hybrid camera configuration. *IEEE Transactions on Robotics*, 23(1), pages 73–86, 2007.

M. Liu, C. Pradalier, F. Pomerleau, and R. Siegwart. The role of homing in visual topological navigation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Portugal, pages 567–572, Oct 2012.

M. Liu, B.T. Alper, and R. Siegwart. An adaptive descriptor for uncalibrated omnidirectional images - towards scene reconstruction by trifocal tensor. In *IEEE International Conference on Robotics and Automation*, Karlsruhe, Germany, pages 558–563, 2013.

G. López-Nicolás, J. J. Guerrero, and C. Sagüés. Visual control through the trifocal tensor for nonholonomic robots. *Robotics and Autonomous Systems*, 58(2), pages 216-226, 2010.

E. Marchand, F. Chaumette, F. Spindler, and M. Perrier. Controlling an uninstrumented manipulator by visual servoing. *International Journal of Robotics Research*, 21(7), pages 635–648, 2002.

E. Marchand, F. Spindler, and F. Chaumette. ViSP for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics and Automation Magazine*, 12(4), pages 40–52, 2005.

E. Montijano, J. Thunberg, X. Hu, and C. Sagüés. Epipolar visual servoing for multirobot distributed consensus. *IEEE Transactions on Robotics*, 29(5), pages 1212–1225, 2013.

Y.K. Nahmias and D.J. Oddel. Analysis of radiation forces in laser trapping and laser-guided direct writing applications. *IEEE Journal of Quantum Electronics*, 38(2), pages 131–141, 2002.

F. Paccot, N. N Andreff, and P. Martinet. A review on the dynamic control of parallel kinematic machines: Theory and experiments. *International Journal of Robotics Research*, 28(3), pages 395–416, 2009.

L. Pari, J.M. Sebastian, A. Traslosheros, and L. Angel. A comparative study between analytic and estimated image jacobian by using a stereoscopic system of cameras. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, pages 6208–6215, 2010.

B. Petrak, K. Konthasinghe, S. Perez, and A. Müller. Feedback-controlled laser fabrication of micromirror substrates. *Review of Scientific Instruments*, 82:123112–123112–6, 2011.

K. Rabenorosoa, B. Tasca, A. Zerbib, T.E. Pengwang, P. Rougeot, and N. Andreff. Squipabot: a mesoscale parallel robot for a laser phonosurgery. In *International Symposium on Optomechatronic Technologies*, pages 1–6, Seattle, USA, 2014.

S. Rajesh and Y. Bellouard. Towards fast femtosecond laser micromachining of glass, effect of deposited energy. In *Conference on Lasers and Electro-Optics and Quantum Electronics and Laser Science Conference* , pages 1–2, 2010.

P. Rives. Visual servoing based on epipolar geometry. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Takamatsu, Japan, pages 602–607, 2000.

B. Rosa, M.S. Erden, T. Vercauteren, B. Herman, J. Szewczyk, and G. Morel. Building large mosaics of confocal endomicroscopic images using visual servoing. *IEEE Transactions on Biomedical Engineering*, 60(4), pages 1041-1049, 2013.

A. Ruf and R. Horaud. Visual servoing of robot manipulators part i: Projective kinematics. *International Journal of Robotics Research*, 18(11), pages 1101–1118, 1999.

D. Sabatta and R. Siegwart. Vision-based path following using the 1d trifocal tensor. In *IEEE International Conference on Robotics and Automation*, Karlsruhe, Germany, pages 3095–3102, 2013.

A. C. Sanderson and L. E. Weiss. *Adaptive visual servo control of robots*. Springer, 1983.

A. Schoob, D. Kundrat, L.A. Kahrs, and T Ortmaier. Real-time surface reconstruction and motion tracking for adapting a preplanned laser incision path in phonomicrosurgery. In *International Conference of the Society for Medical Innovation and Technology*, pages 1–4, 2013.

A. Schoob, D. Kundrat, L.A. Kahrs, L. Kleingrothe, T. Ortmaier, and N. Andreff. Tissue surface information for intraoperative incision planning and focus adjustment in laser surgery. International Journal of Computer Assisted Radiology and Surgery, 10(0), pages 171-181, 2014.

J.A. Séon, B. Tamadazte, N. Andreff. Path following: from mobile robotics to laser surgery. In *3rd Workshop on Visual Control of Mobile Robots (ViCoMoR) at IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, USA, 2014.

J.A. Séon, B. Tamadazte, N. Andreff. Decoupling Path Following and Velocity Profile in Vision-Guided Laser Steering. IEEE Transactions on Robotics, PP(99), doi: 10.1109/TRO.2015.2400660, pages 1-10, 2015.

A. Shademan and M. Jägersand. Three-view uncalibrated visual servoing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, pages 6234–6239, 2010.

R. Shimokita, S.-I. Okihara, S. Eura, and H. Eda. Evaluation of bone carbonization by co2 laser osteotomy. In *World Automation Congress*, pages 1–6, 2008.

M. Shoham and Y. Koren. Elimination of the joint feedback in sensor-guided robots. In *ASME Winter Annual Meeting*, 11, pages 409–418, 1988.

T. Sibillano, A. Ancona, V. Berardi, and P.M. Lugarà. A real-time spectroscopic sensor for monitoring laser welding processes. *Sensors*, 9, pages 3376–3385, 2009.

B. Tamadazte and N. Andreff. Weakly calibrated stereoscopic visual servoing for laser steering: Application to phonomicrosurgery. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Hong Kong, pages 743–748, 2014.

B. Tamadazte, E. Marchand, P. Piat, and S. Dembélé. CAD model based tracking and 3d visual-based control for MEMS microassembly. *International Journal of Robotics Research*, 29(11), pages 1416–1434, 2010.

B. Tamadazte, N. Piat, and E. Marchand. A direct visual servoing scheme for automatic nanopositioning. *IEEE/ASME Transactions on Mechatronics*, 17(4), pages 728–736, 2012.

R. Tatsambon Fomena, F. Chaumette. Improvements on visual servoing from spherical targets using a spherical projection model. *IEEE Transactions on Robotics*, 25(4), pages 874-886, 2009.

F. Wang, J. Cui, S.K. Phang, B.M. Chen, and T.H. Lee. A mono-camera and scanning laser range finder based UAV indoor navigation system. In *International Conference on Unmanned Aircraft Systems*, Atlanta, USA, pages 694–701, 2013.

W.F. Xie, Z. Li, X.W. Tu, and C. Perron. Switching control of image-based visual servoing with laser pointer in robotic manufacturing systems. *IEEE Transactions on Industrial Electronics*, 56(2), pages 520–529, 2009.

X. Yang, P. Jun, Y. Wentao, F. Yuan, and L. Weirong. Three-step epipolar-based visual servoing for nonholonomic robot with fov constraint. *Journal of Applied Mathematics*, pages 1–14, 2014.

N. Zhenjiang, A. Bolopion, J. Agnus, R. Benosman, and S. Régnier. Asynchronous event-based visual shape tracking for stable haptic feedback in microrobotics. *IEEE Transactions on Robotics*, 28(5):1081–1089, Oct 2012.