Preliminary variation on multiview geometry for vision-guided laser surgery.

Nicolas Andreff, Brahim Tamadazte, Sounkalo Dembélé, Zill-e Hussnain
Institut FEMTO-ST, Université de Franche-Comté/CNRS/ENSMM/UTBM,
24 rue Savary, 25000 Besançon, France
Nicolas.Andreff@femto-st.fr

June 14, 2013

Abstract

This paper proposes to use the multiview geometry to control an orientable laser beam for surgery. Two methods are proposed based on the analogy between a scanning laser beam and a camera: the first method uses one camera and the laser scanner as a virtual camera to form a virtual stereoscopic system while the second method uses two cameras to form a virtual trifocal system. Using the associated epipolar or trifocal geometry, two control laws are derived without any matrix inversion nor estimation of the 3D scene. It is shown that the more geometry is used, the simpler the control gets. These control laws show, as expected, exponential convergence in simulation validation.

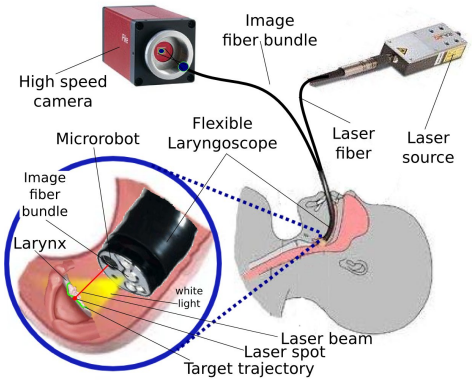

Figure 1: Endoscopic laser microphonosurgery.

1 INTRODUCTION

1.1 Application context

The μRALP project involves the development of a system for endoluminal laser phonosurgery, *i.e.* surgery of the vocal chords using a laser emitted from inside the larynx (Fig. 1). During ablation and even more during resection, the tissues move and change. Moreover, in case of an endoscopic laser steering system, one can not guarantee any time stability of the microrobot calibration. Therefore, visual feedback is crucial to accurate operation.

In addition, it shall be kept in mind that 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, which is to be avoided. As a consequence, laser surgery implies high bandwidth sensing devices and control laws, namely high-speed visual servoing.

Visually-guided laser surgery is not restricted to phonosurgery and can be used in several other medical applications such as otology, dermatology, orthoped or ophtalmology, to name a few. It should gain even higher attention in the future since single-port surgery [Sánchez et al., 2011] and NOTES [Bardou et al., 2009] requires tool minimization. As a consequence, the usual mechanical or electrical scalpels might not be usable anymore because they will not stand the mechanical efforts

imposed by the tissues, on the contrary to the contactless optical scalpel formed by the laser.

1.2 Contribution of the paper

The contribution of this paper is to discuss the control of a laser (namely, the invisible incision laser for surgery intervention with co-axial visible Helium-Neon (HeNe) laser pointer) over any surface using visual feedback. It shows that making call to multiview geometry can simplify the control: no matrix inversion, no explicit knowledge or reconstruction of the 3D scene.

It is developed here in two versions (monocular or stereoscopic observation of the laser) in the so-called eye-to-hand configuration [Chaumette and Hutchinson, 2007] where the laser moves while the camera(s) remain(s) static.

1.3 State of the art

This contribution is thus related to the use of visual servoing using laser(s), stereoscopic visual servoing using the epipolar constraint and visual servoing exploiting the trifocal constraint. Visible lasers have been used in visual servoing for a long time, either in a eye-to-laser configuration [Khadraoui et al., 1996, Amin-Nejad et al., 2003, Krupa et al., 2003, Lv et al., 2010] or in a laser-in-eye one [Andreff et al., 2002, Ginhoux et al., 2004, Xie et al., 2009], according to whether the laser is fixed with respect to the camera or not. In all cases, the laser was essentially used as an additional device to get an estimation on the depth of feature points, not as the *unique* controlled feature itself as we do here. Most of them did not explicitly take into account the epipolar constraint (applicable to the laser+camera system) at the control level, as it was done in many other pieces of work on stereoscopic visual servoing [Hager, 1997, Hespanha et al., 1998, Ruf and Horaud, 1999, Lamiroy et al., 2000, Ge and Jie, 2007, Pari et al., 2010, Alkhalil and Doignon, 2012] or in monocular visual servoing where the epipolar constraint is taken between the initial and desired images [Basri et al., 1999, Rives, 2000, Mariottini et al., 2007].

Going further in the coupling between multiview geometry and visual servoing, some authors have used the trifocal constraint between two successive stereoscopic pairs [Shademan and Jagersand, 2010] or between the initial, desired and current monocular view [López-Nicolás et al., 2010].

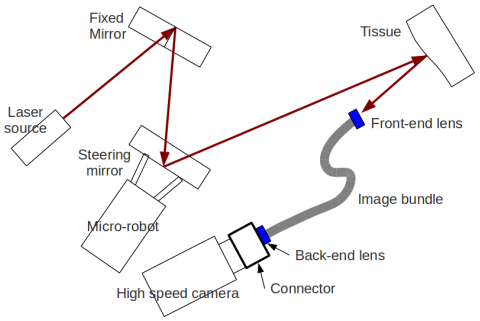

Figure 2: Schematic view of the laser steering system with one camera.

2 Monocular visually-guided laser surgery

In this section, we analyze the control of the laser spot with the microrobot using an optic fiber bundle to bring the image of the scene onto a high-speed camera (Fig. 1 and 2).

This control can be done in two ways: using the standard visual servoing equations or using the above grounding analogy.

2.1 A word on standard control

Let us note $\underline{\boldsymbol{z}}$ the direction of the laser beam reflecting from the steering mirror towards the surface, $\boldsymbol{P}$ the position of the laser spot on the surface, and $\boldsymbol{p}$ the position of the laser spot in the image. Then, it is trivial to write, in the reference frame R_0 attached to the zero-reference (rotating center) of the steering mirror:

$$^0\boldsymbol{P} = d\,^0\underline{\boldsymbol{z}} \tag{1}$$

where d is the distance traveled by the laser from the mirror to the surface.

This distance can not be measure, to the contrary of $^0\underline{\boldsymbol{z}}$ which can be obtained from the microrobot encoders. However, it can be modeled if one approximates the surface in $\boldsymbol{P}$ by a plane of equation:

$$^0\underline{\boldsymbol{n}}^{T\,0}\boldsymbol{P} - d_0 = 0 \tag{2}$$

where ${}^0\underline{\boldsymbol{n}}$ is the orientation of the surface normal in R_0 and d_0 is the distance of the plane to the origin of R_0. Using this model, one finds:

$$d = \frac{d_0}{{}^0\underline{\boldsymbol{n}}^T\,{}^0\underline{\boldsymbol{z}}} \tag{3}$$

2.1.1 Standard derivation

On the other hand, the perspective projection equation yields:

$$\tilde{\boldsymbol{p}} = \mathbf{K}\frac{{}^c\boldsymbol{P}}{Z} \tag{4}$$

where $\mathbf{K}$ is the matrix containing the intrinsic parameters, $\boldsymbol{P}$ is now expressed in the camera frame R_c, Z is the unmeasured depth along the line of sight passing through $\boldsymbol{p}$ as well as the third coordinate of ${}^c\boldsymbol{P}$ and $\tilde{\boldsymbol{p}}$ represents the homogeneous coordinates of $\boldsymbol{p}$.

To apply, the usual visual servoing approach, one needs to differentiate the latter with time:

$$\dot{\tilde{\boldsymbol{p}}} = \frac{1}{Z}\mathbf{K}\begin{pmatrix}\mathbf{I}_{2\times2} & -\boldsymbol{p}\\ \mathbf{0}_{1\times2} & 0\end{pmatrix}{}^c\dot{\boldsymbol{P}} \tag{5}$$

One has another expression for ${}^c\dot{\boldsymbol{P}}$ by differentiating (1):

$${}^0\dot{\boldsymbol{P}} = \dot{d}\,{}^0\underline{\boldsymbol{z}} + d\,{}^0\dot{\underline{\boldsymbol{z}}} \tag{6}$$

and expressing the latter in R_c:

$${}^c\dot{\boldsymbol{P}} = {}^c\mathbf{R}_0\left(\dot{d}\,{}^0\underline{\boldsymbol{z}} + d\,{}^0\dot{\underline{\boldsymbol{z}}}\right) \tag{7}$$

Now, from 3, one gets (under the simplifying assumption that the surface plane does not change):

$$\dot{d} = -\frac{d_0\,{}^0\underline{\boldsymbol{n}}^T}{\left({}^0\underline{\boldsymbol{n}}^T\,{}^0\underline{\boldsymbol{z}}\right)^2}\,{}^0\dot{\underline{\boldsymbol{z}}} \tag{8}$$

Putting (7) and (8) into (5) gives therefore $\dot{\tilde{\boldsymbol{p}}}$ under the form:

$$\dot{\tilde{\boldsymbol{p}}} = \mathbf{L}(d, Z, {}^0\underline{\boldsymbol{n}}, d_0, {}^c\mathbf{R}_0, \boldsymbol{p}, {}^0\underline{\boldsymbol{z}})\,{}^0\dot{\underline{\boldsymbol{z}}} \tag{9}$$

where

$$\mathbf{L} = \frac{d}{Z}\mathbf{K}\begin{pmatrix}\mathbf{I}_{2\times2} & -\boldsymbol{p}\\ \mathbf{0}_{1\times2} & 0\end{pmatrix}{}^c\mathbf{R}_0\left(\mathbf{I}_3 - \frac{{}^0\underline{\boldsymbol{z}}\,{}^0\underline{\boldsymbol{n}}^\top}{{}^0\underline{\boldsymbol{n}}^\top\,{}^0\underline{\boldsymbol{z}}}\right) \tag{10}$$

is of dimension 3×3 but has only rank 2.

Pseudo-inverting it allows converting the image velocity of the laser spot into the velocity of the laser beam, which, in turn, shall be converted into microrobot velocity through the differential inverse kinematic model.

Figure 3: The standard controller shows a convergence to the desired position despite a non-straight trajectory.

Now, we can come up to the control law, by enforcing a first order behaviour of the error in the image between the current and the desired projections of the laser spot:

$$\dot{\tilde{\boldsymbol{p}}} = -\lambda(\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{p}}^*) \tag{11}$$

or, if one wishes to track a trajectory:

$$\dot{\tilde{\boldsymbol{p}}} = -\lambda(\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{p}}^*) - \dot{\tilde{\boldsymbol{p}}}^* \tag{12}$$

This control law has been tested and shows a convergence to the desired position (global minimum) despite a non-straight trajectory (see Fig. 3). The reason for this is that the numerical pseudo-inversion does not take into account geometry. Therefore, a better solution is as follows.

2.1.2 Alternate derivation

Firstly, one cans pseudo-invert (5) "at hand" as:

$${}^c\dot{\boldsymbol{P}} = Z\mathbf{K}^{-1}\dot{\tilde{\boldsymbol{p}}} + \mu\mathbf{K}^{-1}\tilde{\boldsymbol{p}} \tag{13}$$

where $\mathbf{K}^{-1}\tilde{\boldsymbol{p}}$ is the kernel of (5) and μ is a scalar. The latter can be defined, using the time derivative of (2) to get a constraint on the velocity of $\boldsymbol{P}$. After a couple of algebraic manipulations, one gets:

$$\mu = Z\frac{{}^0\underline{\boldsymbol{n}}^\top\,{}^0\mathbf{R}_c\mathbf{K}^{-1}\dot{\tilde{\boldsymbol{p}}}}{{}^0\underline{\boldsymbol{n}}^\top\,{}^0\mathbf{R}_c\mathbf{K}^{-1}\tilde{\boldsymbol{p}}} \tag{14}$$

from which, we deduce the final input-output relationship from $\dot{\tilde{\boldsymbol{p}}}$ to ${}^0\dot{\underline{\boldsymbol{z}}}$:

$${}^0\dot{\underline{\boldsymbol{z}}} = \frac{Z}{d}\left(\mathbf{I}_3 - {}^0\underline{\boldsymbol{z}}\,{}^0\underline{\boldsymbol{z}}^\top\right)\left(\mathbf{I}_3 - \frac{{}^0\boldsymbol{v}\,{}^0\underline{\boldsymbol{n}}^\top}{{}^0\underline{\boldsymbol{n}}^\top\,{}^0\boldsymbol{v}}\right){}^0\mathbf{R}_c\mathbf{K}^{-1}\dot{\tilde{\boldsymbol{p}}} \tag{15}$$

with ${}^0\boldsymbol{v} = {}^0\mathbf{R}_c\mathbf{K}^{-1}\tilde{\boldsymbol{p}}$ and the controller by using (11) or (12).

Therefore, this controller needs the estimation of d and Z (which can be taken as constant since they act as a gain), but more crucially the robot-camera calibration (${}^0\mathbf{R}_c$), the camera calibration ($\mathbf{K}$) and the scene structure (${}^0\underline{\boldsymbol{n}}$).

However, it has the following advantages over the standard visual servoing approach:

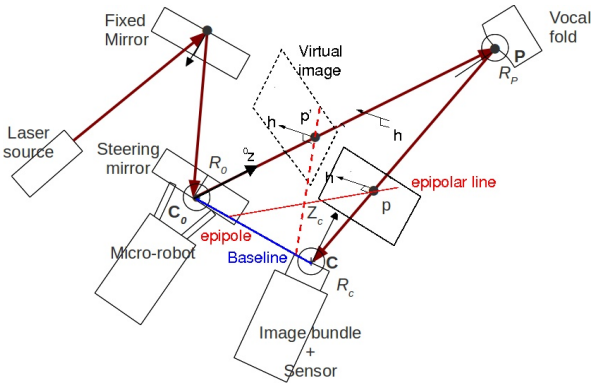

Figure 4: Analogy with stereoscopy.

- it does not involve any matrix inversion;
- it can work without any 3D reconstruction.

2.2 Laser visual servoing using epipolar geometry

On the opposite to the above method, which is totally generic, the method we propose here is totally hand-made and tailored to the specific case of a laser beam being observed by a camera.

Indeed, the set-up seen in Fig. 2 is analogous to a degenerate case of epipolar geometry in Fig. 4. Thereby, points $\boldsymbol{p}$ and $\boldsymbol{p}'$ are the images of the same spatial point $\boldsymbol{P}$ and are hence linked by the epipolar constraint:

$$\tilde{\boldsymbol{p}}'^{\top}\mathbf{F}\tilde{\boldsymbol{p}} = 0 \tag{16}$$

where $\mathbf{F}$ is the fundamental matrix of the 2-views system [Hartley and Zisserman, 2000]. Actually, this epipolar constraint is defined up to a scale factor, and thereby, $\tilde{\boldsymbol{p}}'$ can be replaced by ${}^0\underline{\boldsymbol{z}}$, the unit vector describing, in the microrobot base frame R_0, the direction of the laser beam from the mirror to the surface:

$${}^0\underline{\boldsymbol{z}}^{\top}\mathbf{F}\tilde{\boldsymbol{p}} = 0 \tag{17}$$

This equation expresses the fact that the origin of the camera, the pivot point of the microrobot (rotating center), the laser beam, the line of sight and the laser spot on the surface are coplanar. It can be also interpreted in three ways:

$${}^0\underline{\boldsymbol{z}} \perp \mathbf{F}\tilde{\boldsymbol{p}} \tag{18}$$

$$\tilde{\boldsymbol{p}} \perp \mathbf{F}^{\top}\,{}^0\underline{\boldsymbol{z}} \tag{19}$$

and both $\mathbf{F}\tilde{\boldsymbol{p}}$ and $\mathbf{F}^{\top}\,{}^0\underline{\boldsymbol{z}}$ represent the (non-unit) normal vector to the epipolar plane in the microrobot and the camera frame, respectively.

The time derivative of the epipolar constraint is:

$$({\mathbf{F}^{\top}\,}^0\underline{\boldsymbol{z}})^{\top}\dot{\tilde{\boldsymbol{p}}} + (\mathbf{F}\tilde{\boldsymbol{p}})^{\top}\,{}^0\dot{\underline{\boldsymbol{z}}} = 0 \tag{20}$$

Now, we can decompose ${}^0\dot{\underline{\boldsymbol{z}}}$ into a component orthogonal to the epipolar plane and a component inside the latter:

$${}^0\dot{\underline{\boldsymbol{z}}} = \alpha\,{}^0\underline{\boldsymbol{h}} + \beta\,{}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{h}} \tag{21}$$

where ${}^0\underline{\boldsymbol{h}} = \frac{\mathbf{F}\tilde{\boldsymbol{p}}}{\|\mathbf{F}\tilde{\boldsymbol{p}}\|}$. Replacing this expression into the epipolar constraint and reordering the terms, we get:

$$\alpha = -\frac{({\mathbf{F}^{\top}\,}^0\underline{\boldsymbol{z}})^{\top}}{\|\mathbf{F}\tilde{\boldsymbol{p}}\|}\dot{\tilde{\boldsymbol{p}}} \tag{22}$$

Actually, α only depends on the projection of $\dot{\tilde{\boldsymbol{p}}}$ onto the normal to the epipolar plane, but expressed in the camera frame, *i.e.* along ${}^c\underline{\boldsymbol{h}} = \frac{\mathbf{F}^{\top}\,{}^0\underline{\boldsymbol{z}}}{\|\mathbf{F}^{\top}\,{}^0\underline{\boldsymbol{z}}\|}$. Thus, the remaining part of $\dot{\tilde{\boldsymbol{p}}}$ is obtained by canceling this projection:

$$\dot{\tilde{\boldsymbol{p}}} = a\,{}^c\underline{\boldsymbol{h}} + (\mathbf{I}_3 - {}^c\underline{\boldsymbol{h}}\,{}^c\underline{\boldsymbol{h}}^{\top})\dot{\tilde{\boldsymbol{p}}} \tag{23}$$

where the value of a does not have any interest for the sequel, but can be related to α by inserting the latter equation into the former.

Now, concentrate on the part of $\dot{\tilde{\boldsymbol{p}}}$ lying in the epipolar plane. In (23), it is expressed in the camera frame, so we just need to bring it back to the microrobot frame, going backwards the camera intrinsic parameters, the orientation of the camera frame with respect to the microrobot frame and compensating for the unknown scale factor in $\mathbf{F}$, to get β:

$$\beta = \frac{\|\mathbf{F}^{\top}\,{}^0\underline{\boldsymbol{z}}\|}{\|\mathbf{F}\tilde{\boldsymbol{p}}\|}({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{h}})^{\top}\,{}^0\mathbf{R}_c\mathbf{K}^{-1}(\mathbf{I}_3 - {}^c\underline{\boldsymbol{h}}\,{}^c\underline{\boldsymbol{h}}^{\top})\dot{\tilde{\boldsymbol{p}}} \tag{24}$$

As a consequence, we have expressed ${}^0\dot{\underline{\boldsymbol{z}}}$ as a function of $\dot{\tilde{\boldsymbol{p}}}$:

$${}^0\dot{\underline{\boldsymbol{z}}} = \{({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{h}})({}^0\underline{\boldsymbol{z}} \times {}^0\underline{\boldsymbol{h}})^{\top}\,{}^0\mathbf{R}_c\mathbf{K}^{-1}(\mathbf{I}_3 - {}^c\underline{\boldsymbol{h}}\,{}^c\underline{\boldsymbol{h}}^{\top}) - {}^0\underline{\boldsymbol{h}}\,{}^c\underline{\boldsymbol{h}}^{\top}\}\frac{\|\mathbf{F}^{\top}\,{}^0\underline{\boldsymbol{z}}\|}{\|\mathbf{F}\tilde{\boldsymbol{p}}\|}\dot{\tilde{\boldsymbol{p}}} \tag{25}$$

Consequently, we have the exact expression of the conversion of the image velocity into the laser

beam velocity, without any matrix inversion, nor any explicit triangulation or scene structure knowledge. This expression only depends on the measurements ($^0\underline{\boldsymbol{z}}$ and $\boldsymbol{p}$), the fundamental matrix $\mathbf{F}$ and a reduced set of calibration parameters ($\mathbf{K}$ and $^c\mathbf{R}_0$).

Finally, $^0\underline{\boldsymbol{z}}$ is a rotating unit vector according to the microrobot pan-and-tilt velocities $\boldsymbol{\omega}$ and hence:

$$^0\dot{\underline{\boldsymbol{z}}} = 2\boldsymbol{\omega} \times {}^0\underline{\boldsymbol{z}} \tag{26}$$

where the factor 2 comes from the reflection law in optics. Thereby, we get the control to send to the microrobot:

$$\boldsymbol{\omega} = \frac{1}{2}\, {}^0\underline{\boldsymbol{z}} \times {}^0\dot{\underline{\boldsymbol{z}}} \tag{27}$$

where $^0\dot{\underline{\boldsymbol{z}}}$ is obtained from (25) and either (11) or (12).

2.3 Advantages

The control law based on epipolar geometry has the following advantages over the standard visual servoing approach:

- it does not involve any matrix inversion;
- it does not require any 3D reconstruction;
- it does not require any scene structure prior knowledge.

3 Stereoscopic visually-guided laser surgery

Now, we take into account two cameras observing the surface: two optic fiber bundles bring the two images of the scene onto a high-speed camera 5).

3.1 Trifocal constraint

Now, we have three epipolar constraints:

$$\tilde{\boldsymbol{p}}_L{}^\top\, {}^L\mathbf{F}_R\, \tilde{\boldsymbol{p}}_R = 0 \tag{28}$$

$$^0\underline{\boldsymbol{z}}^\top\, {}^0\mathbf{F}_R\, \tilde{\boldsymbol{p}}_R = 0 \tag{29}$$

$$^0\underline{\boldsymbol{z}}^\top\, {}^0\mathbf{F}_L\, \tilde{\boldsymbol{p}}_L = 0 \tag{30}$$

From the last two ones, we can deduce that $^0\underline{\boldsymbol{z}}$ is orthogonal to the epipolar planes made by the laser

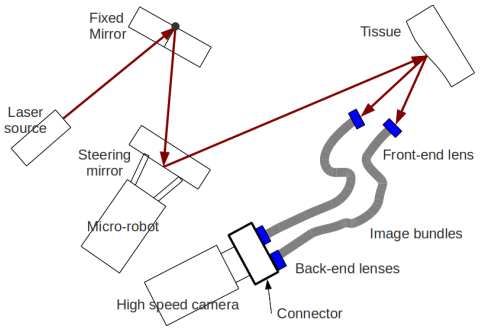

Figure 5: Schematic view of the laser steering system with two cameras.

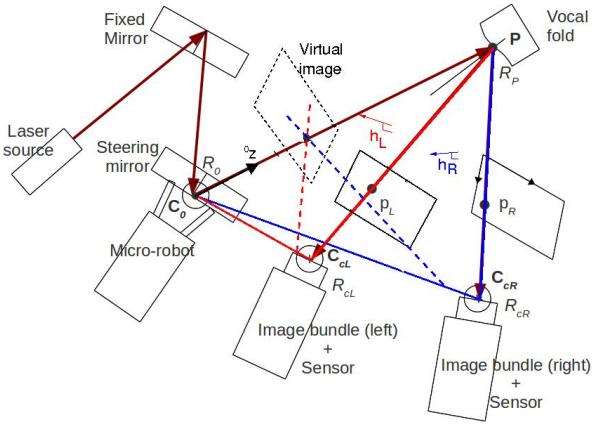

Figure 6: Analogy with trifocal geometry.

beam and, respectively, the left and right image points:

$$\boldsymbol{h}_R = {}^0\mathbf{F}_R \tilde{\boldsymbol{p}}_R \quad \boldsymbol{h}_R = {}^0\mathbf{F}_L \tilde{\boldsymbol{p}}_L \tag{31}$$

Therefore, it is parallel to their cross product, which can be expressed using the "cross-product trick":

$${}^0\underline{\boldsymbol{z}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) = 0 \tag{32}$$

which is nothing but the trifocal constraint (maybe expressed in a simpler manner).

3.2 Control

Let us differentiate the above trifocal constraint with time:

$$\begin{aligned} &{}^0\dot{\underline{\boldsymbol{z}}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) + {}^0\underline{\boldsymbol{z}} \times \left(({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) \times \boldsymbol{h}_L \right) \\ &+ {}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) = 0 \end{aligned} \tag{33}$$

which reorganizes as:

$$\begin{aligned} &{}^0\dot{\underline{\boldsymbol{z}}} \times (\boldsymbol{h}_R \times \boldsymbol{h}_L) = \\ &\quad {}^0\underline{\boldsymbol{z}} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \end{aligned} \tag{34}$$

From the trifocal constraint, we also get trivially:

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

and hence

$$\begin{aligned} &{}^0\dot{\underline{\boldsymbol{z}}} \times {}^0\underline{\boldsymbol{z}} = \\ &\quad \frac{{}^0\underline{\boldsymbol{z}}}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \end{aligned} \tag{36}$$

where one recognizes $\boldsymbol{\omega}$ from (27) on the first line:

$$\boldsymbol{\omega} = \frac{1}{2} \frac{{}^0\underline{\boldsymbol{z}}}{\|\boldsymbol{h}_R \times \boldsymbol{h}_L\|} \times \left(\boldsymbol{h}_L \times ({}^0\mathbf{F}_R \dot{\tilde{\boldsymbol{p}}}_R) - \boldsymbol{h}_R \times ({}^0\mathbf{F}_L \dot{\tilde{\boldsymbol{p}}}_L) \right) \tag{37}$$

To finalize the control, one just needs to set the desired control in each images, as in (11) or (12).

3.3 Advantages

The "trifocal" control law has several advantages over the other two ones:

- it does not involve any matrix inversion;
- it does not make any call to any explicit 3D reconstruction;
- it does not require any prior knowledge on the scene structure;
- it does not need any proprioceptive sensing of the micro-mirror configuration;
- an admissible control reference can be directly recorded by the surgeon directly in the image pair;
- it only requires weak calibration (Fl and Fr) of the system;
- it is so simple that global proof of stability seems into reach.

4 Simulation results

4.1 Without multiview geometry

The alternate solution to the standard controller was implemented in a simulator and yields exponential decay of the image errors (Fig. 7). It can be seen that it is fully decoupled: the image trajectory is a straight line.

4.2 Use of epipolar geometry

The control law based on epipolar geometry was implemented in a simulator and gives exponential decay of the image errors (Fig. 8). The perfect decoupling is lost, but there is no need any more to know the scene structure.

4.3 Use of trifocal constraint

The control based on trifocal constraint was implemented in a simulator and gives exponentially decay of the image errors (Fig.9). It can be noticed that both image trajectories are straight (perfect decoupling) and that without knowledge on the scene structure.

In the future, all the proposed control laws will be tested further with the introduction of noise in the sensor signals and the calibration parameters. They are being implemented on a basic hardware (standard camera + pan/tilt micropositionner from Physical Instruments Inc. (PI)) before it is implemented in an endoscopic set-up.

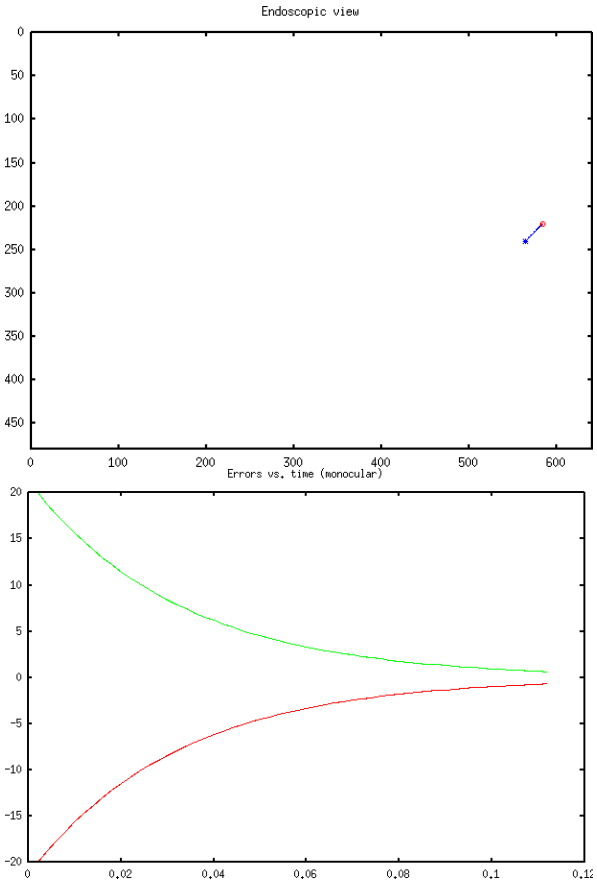

Figure 7: Image trajectory and time evolution of the image errors in the case of monocular visually guided laser surgery (without multiview geometry).

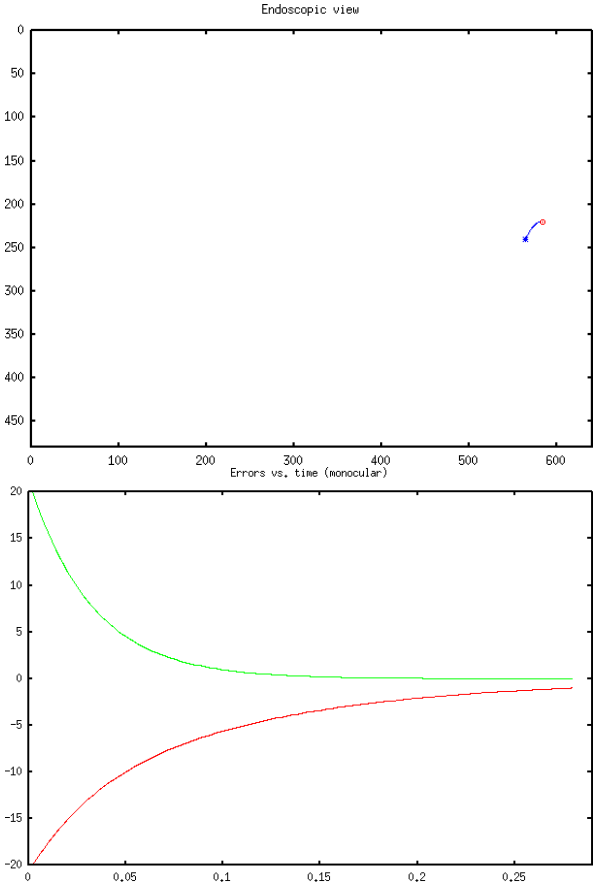

Figure 8: Image trajectory and time evolution of the image errors in the case of monocular visually guided laser surgery (epipolar geometry).

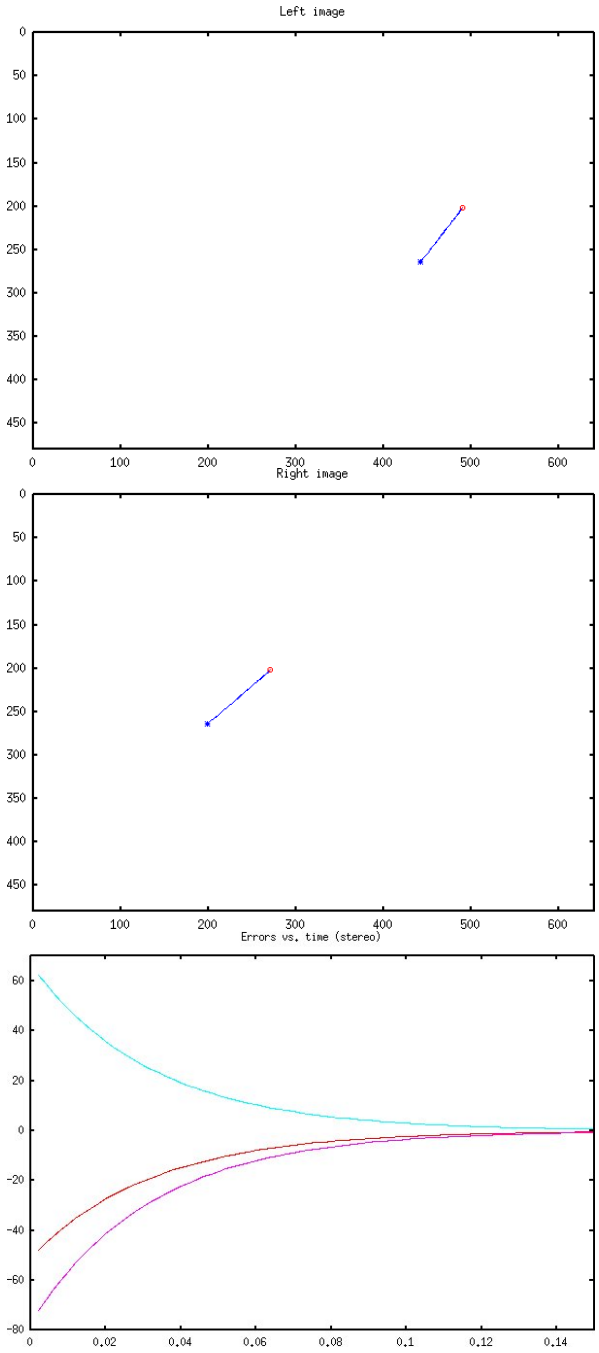

Figure 9: Image trajectories and time evolution of the image errors in the case of stereoscopic visually guided laser surgery (trifocal constraint).

5 Discussion

In this paper, it was shown that resorting to geometry simplifies the eye-to-hand control law for a surgical laser (and any other application where a beam needs be accurately swept over a surface). The further we dig into geometry, the simpler the control law becomes.

Indeed, using epipolar geometry allows for relieving us from estimating the normal to the surface, which is needed by the "plain" visual servoing control law. It also allows for an implicit rather than explicit triangulation. Yet, the desired reference point or trajectory in the image needs be admissible, *i.e.* coherent with the surface geometry and relative positioning with respect to the micro-mirror. This, if a dedicated surgeon-robot interface is designed [Mattos et al., 2011] to define the desired trajectory in the image, then the latter will geometrically contain a coherent description of the 3D surface, and thus, the 3D information is not purely and simply thrown away as it could seem but, rather, it is implicitly used.

Further investigation of geometry, namely the trifocal geometry associated to a stereoscopic observation of the laser spot, was shown to further simplify the control. One does not need any strong calibration anymore. The definition of the desired trajectory in the pair of images makes it more reliable as far as its coherence with the surface geometry is concerned. This can be obtained, for instance, by extending to 3D, the dedicated surgeon-robot interface in [Mattos et al., 2011], which in turn could be done by the current 2D set-up complemented with epipolar-based matching. The use of a pair of images should also increase the resolution of the spot positioning, namely by using the optimal stereoscopic point detection proposed by Hartley and Zissermann [Hartley and Zisserman, 2000]. It is thus expected to allow for a simple automated calibration procedure and for an increase of the control robustness, which are key issues in the transfer of automation into actual clinical devices.

Also, many micromanipulators have a parallel kinematics architecture, which are known to be controllable without any joint sensing [Andreff and Martinet, 2009]. Consequently, using the proposed multiview geometric approach might enable simplified miniaturization of laser steering in an endoscopic set-up, because one

can design steering parallel kinematics mechanisms without proprioceptive sensors. This is a very crucial investigation field, since endoscopic laser surgery faces very contradictory requirements in terms of sweeping range and frequency (yielding larger mechanisms) and of available space at the endoscopic tip.

ACKNOWLEDGEMENTS

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`).

References

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

[Amin-Nejad et al., 2003] Amin-Nejad, S., Smith, J., and Lucas, J. (2003). A visual servoing system for edge trimming of fabric embroideries by laser. *Mechatronics*, 13(6):533–551.

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

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

[Bardou et al., 2009] Bardou, B., Nageotte, F., Zanne, P., and de Mathelin, M. (2009). Design of a telemanipulated system for transluminal surgery‡. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 5577–5582. IEEE.

[Basri et al., 1999] Basri, R., Rivlin, E., and Shimshoni, I. (1999). Visual homing: Surfing on the epipoles. *International Journal of Computer Vision*, 33(2):117–137.

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

[Ge and Jie, 2007] Ge, L. and Jie, Z. (2007). A real-time stereo visual servoing for moving object grasping based parallel algorithms. In *Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference on*, pages 2886–2891.

[Ginhoux et al., 2004] Ginhoux, R., Gangloff, J., de Mathelin, M., Soler, L., Sanchez, M. A., and Marescaux, J. (2004). Beating heart tracking in robotic surgery using 500 hz visual servoing, model predictive control and an adaptive observer. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, volume 1, pages 274–279. IEEE.

[Hager, 1997] Hager, G. D. (1997). A modular system for robust positioning using feedback from stereo vision. *IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION*, 13(4).

[Hartley and Zisserman, 2000] Hartley, R. and Zisserman, A. (2000). *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press.

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

[Khadraoui et al., 1996] Khadraoui, D., Motyl, G., Martinet, P., Gallice, J., and Chaumette, F. (1996). Visual servoing in robotics scheme using a camera/laser-stripe sensor. *Robotics and Automation, IEEE Transactions on*, 12(5):743–750.

[Krupa et al., 2003] Krupa, A., Gangloff, J., Doignon, C., de Mathelin, M., Morel, G., Leroy, J., Soler, L., and Marescaux, J. (2003). Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE Transactions on Robotics and Automation*, 19(5).

[Lamiroy et al., 2000] Lamiroy, B., Espiau, B., Andreff, N., and Horaud, R. (2000). Controlling robots with two cameras: How to do it properly. In *IEEE Int. Conf. on Robotics and Automation (ICRA'00)*, San Francisco, California, USA.

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

[Lv et al., 2010] Lv, S., Zhang, D., and Gu, J. (2010). Research of automatic needle locating based on stereo visual servoing. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, volume 4, pages 1779–1783.

[Mariottini et al., 2007] Mariottini, G. L., Oriolo, G., and Prattichizzo, D. (2007). Image-based visual servoing for nonholonomic mobile robots using epipolar geometry. *Robotics, IEEE Transactions on*, 23(1):87–100.

[Mattos et al., 2011] Mattos, L. S., Dagnino, G., Becattini, G., Dellepiane, M., and Caldwell, D. G. (2011). A virtual scalpel system for computer-assisted laser microsurgery. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 1359–1365. IEEE.

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

[Rives, 2000] Rives, P. (2000). Visual servoing based on epipolar geometry. In *Intelligent Robots and Systems, 2000.(IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on*, volume 1, pages 602–607. IEEE.

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

[Sánchez et al., 2011] Sánchez, L. A., Petroni, G., Piccigallo, M., Scarfogliero, U., Niccolini, M., Liu, C., Stefanini, C., Zemiti, N., Menciassi, A., Poignet, P., et al. (2011). Real-time control and evaluation of a teleoperated miniature arm for single port laparoscopy. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 7049–7053. IEEE.

[Shademan and Jagersand, 2010] Shademan, A. and Jagersand, M. (2010). Three-view uncalibrated visual servoing. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 6234–6239. IEEE.

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